Fundamental Concepts of Mathematics

K.T. Leung and P.H.Cheung

Fundamental Concepts of Mathematics

FUNDAMENTAL CONCEPTS OF MATHEMATICS

K.T. Leung and P.H. Cheung

Hong Kong University Press
香港大學出版社

Hong Kong University Press
139 Pokfulam Road, Hong Kong

First published 1988
Reprinted 1994

ISBN 962 209 181 4

Printed in Hong Kong by Condor Production Co. Ltd., Hong Kong.

CONTENTS

PREFACE

Fundamental Concepts of Mathematics is the first of several volumes of a proposed series on fundamental mathematics which serve primarily as textbooks for students in preparation for the A-Level or other public examinations. However they go into more depth than what is required by these examinations, and contain topics that are useful but often omitted for lack of time in the undergraduate courses. Therefore the books of this series can also be used by university students as reference books for supplementary reading. Further volumes on topics in algebra, geometry and calculus are forthcoming. It is sincerely hoped that in their entirety they will ease the dire shortage of suitable textbooks for A-Level mathematics that has persisted over decades in Hong Kong.

The present volume gives a straightforward account of the various number systems of fundamental mathematics. Chapter 1 is a brief account of the informal set language. The reader may find this more accessible and easier than the presentation in *Elementary Set Theory* by D..L.C. Chen and the author.

Natural numbers are studied in Chapters 2 and 3. No attempt has been made to define natural numbers in terms of primitive concepts and axioms. Only the principle of mathematical induction is singled out for a detailed scrutiny in Chapter 2. The well-ordering property of natural numbers is assumed, and the principle of mathematical induction is derived and then studied from different points of views. Natural numbers are used in counting and Chapter 3 is devoted to counting the different ways in which an event can take place. A balls-into-boxes model is used to study the various problems of permutations and combinations. This unified approach to the topics is probably new at this level of school mathematics.

Chapter 4 deals with elementary number theory. At present, bits and pieces of this subject are scattered throughout the 13-year curriculum from class primary one to form upper six. No effort has been made anywhere to present it in a systematic manner. It is not unusual to find beginners at the university who could cope with complicated integration but had no idea that the greatest common divisor can be written as a linear combination. The present chapter is an attempt to redress this shortcoming and to introduce the subject up to the unique factorization theorem and the Chinese remainder theorem. Students in the upper-forms will have no difficulty in working through this chapter and undergraduate students may find it useful to read the chapter before they take up a course on algebra or number theory.

The next two chapters on the real numbers are the most difficult part of the book. Chapter 5 brings out the similarities and the essential differences between the rational numbers and the real numbers. It also

leads to the density theorem and the postulate of continuity. Powers, roots and logarithms of real numbers are then given rigorous definitions. The postulate is put to further use in Chapter 6 where the fundamental notion of convergence is discussed. The idea of limit is introduced first informally by a series of definitions and examples. A precise definition is then formulated after the reader is sufficiently familiar with the main theme. Undue formalism and abstraction are avoided, and proofs that demand more maturity are put in the appendix. The material of these two chapters exceeds the requirement of the A-Level examination. Undergraduate students will find in them many useful detailed discussions that may be omitted in the lectures.

Chapter 7 introduces complex numbers as the ultimate result of a series of attempts to remove the inadequacy of the previous number systems in the provision of solutions to polynomial equations. The system of complex numbers is viewed geometrically as a two-dimensional extension of the one-dimensional system of real numbers. The reader may find the approach to the imaginary unit by rotations both natural and interesting. In the concluding sections, the analytic geometry of straight lines and circles appears in a new form as geometry of complex numbers. This may be taken as an illustration of the adaptability of mathematical ideas and of the fact that seemingly different branches of mathematics are actually well connected. The appendix to this chapter outlines the further development of the number concept in the last century. All material of the chapter is within the reach of upper-form students.

Several hundred exercises are included in the book, with the more difficult ones marked by an asterisk. Answers to some are given at the end of the book. It is not necessary that the reader should try each of them. These excellent exercises are provided by my former student and friend, Mr. P.H. Cheung, who has many years of teaching experience. I would like to take this opportunity to express my deep appreciation to him for his contribution. To my colleagues Dr. M.K. Siu and Dr. K.M. Tsang, I am indebted for their suggestions, comments and criticisms.

K.T. Leung

University of Hong Kong
July 1987

1. Set Notation

Every environment breeds its own language and common words tend to acquire different meanings for different groups of people. For example, a Mediterranean farmer, a chairman of a committee and a young lady may have quite different ideas about the word 'date'. A peasant, a physicist, a mathematician and a race course manager may also think of quite different things when they hear about 'fields'. Mathematics has a vocabulary of its own, in which some common words have very special meanings. The purpose of this first chapter is to explain the precise meaning and the proper usage of a small collection of special terms that we shall encounter in the subsequent chapters.

1.1. OBJECTS

The first word that calls for an explanation is the word 'object'. By an *object* we mean an individual thing of the material world or of our intuition and thought. In this book we shall mainly deal with various types of numbers. Numbers and other mathematical entities of our discussion are therefore objects in the above sense of the word.

In a discussion on any subject an object is always referred to by its names. For example, we use a name of a person to talk about that person and a name of a city to talk about that city. Mathematical objects that appear in our discussion will be given names, usually letters, to be referred to. Thus by saying that *'a number is denoted by x'* or *'x is a number'*, we mean that 'x' is a name of the number in question. Moreover, it is also understood that throughout that particular discussion, 'x' shall not be used as a name of any other object. On the other hand, similar to the usual practice of using different names for one and the same person, it may happen that a mathematical object has different names within one and the same discussion. For example, we may have denoted by x the number that yields six when multipled by two, and later in the same discussion, we may have called y the number which yields seven when added to four. In

that discussion, therefore, both 'x' and 'y' are names of the same number: three.

In mathematics, as well as in other sciences, we examine some special individual objects, and compare one object with another object from different points of view, in order to find out useful properties of the objects under study and important relationships between such objects. Since we have adopted the broadest possible meaning for the word 'object', in general we can only compare objects by means of their identities. In other words, given any two objects, they are either identical or distinct (even though, at times, it may be very hard to determine which is the case). Thus we accept

1.1.1 Rule. *Given any objects x and y, either they are identical (we write $x = y$) or they are distinct (we write $x \neq y$).*

The rules for the proper usage of the *equality sign* $=$ are as follows:

1.1.2 Rules *The following statements hold for all objects x, y and z:*

(a) $x = x$ *(the reflexive law of equality)*

(b) if $x = y$, *then* $y = x$ *(the symmetric law of equality)*

(c) if $x = y$ *and* $y = z$, *then* $x = z$ *(the transitive law of equality).*

1.2 SETS

Many of the objects that we shall study are themselves collections of objects. These collections or *sets* may be finite or infinite; later we shall meet sets with additional structure, but for the time being we shall study sets in general.

Any well-defined collection of objects is called a *set*. The objects which make up a set will be called *elements* or *members* of the set. If an object x is an element of a set A then we write

$$x \in A,$$

in which case we say that x *belongs to* A, x *is in* A or A *contains* x; otherwise we write

$$x \notin A.$$

In saying that a set is a well-defined collection of objects, we mean that we shall adopt the following rule for the proper usage of the sign $\in$.

1.2.1. Rule. *Given any object x and any set A, either $x \in A$ or $x \notin A$.*

For example, if x is George Washington, A is the set of all governors of Hong Kong past and present, and B is the set of all presidents of the United States of America past and present, then obviously $x \notin A$ and $x \in B$. On the other hand, it is much harder to determine whether the numbers $2^{44497} - 1$ is an element of the set of all prime numbers.

We can compare sets by means of equality. Thus given two sets A and B, either $A = B$ or $A \neq B$. We now propose to link up equality (=) of sets with membership ($\in$) of sets by the following rule:

1.2.2. The rule of extension. *Two sets A and B are equal if and only if A contains every element of B and B contains every element of A.*

Using the two symbols = and $\in$, we may write the above rule in the following form:

1.2.3. Theorem. *Let A and B be sets. For $A = B$, it is necessary and sufficient that the following two conditions are satisfied:*

(a) for every object x, if $x \in B$ then $x \in A$

(b) for every object x, if $x \in A$ then $x \in B$.

Similar to a grammatical rule of an ordinary language, the rule of extension is a regulation for the proper usage of set notation; it stipulates that a set is completely determined by its elements.

One consequence of the rule of extension is therefore that every complete membership list defines a set. For example, the list 1, 7, 9, 20 of numbers defines a set whose elements are exactly these four numbers; without ambiguity we can denote this set by putting its complete membership list between a pair of braces:

$$\{1, 7, 9, 20\}.$$

Similarly $\{x + y, x^2 + 2x + 1\}$ is the set that consists of exactly two polynomials $x + y$ and $x^2 + 2x + 1$.

We note that in using the braces notation, we may write the elements of the set any number of times and in any order. For example, the set $\{1, 7, 9, 20\}$ and the set $\{9, 20, 7, 1, 7, 9, 20\}$ are equal because the first set contains every element of the second set and the second set contains every element of the first set. Neat and convenient as it is, the braces

notation has a major limitation. It is only appropriate for sets whose membership lists are short or show some easily recognized regular pattern.

1.2.4. *Example.* $\{3\}$ is the set that consists a single element 3.

1.2.5. *Example.* $\{1, 2, 3, \ldots, 100\}$ is the set of the first one hundred positive integers.

1.2.6. *Example.* $\{2, 4, 6, 8, \ldots\}$ is the set of all positive even integers.

Finally let us pay special attention to sets with very short membership lists. Given any object a, the set that consists of a alone is denoted by $\{a\}$.

A set of one element is called a *singleton set* or simply a *singleton*. For singleton sets we obviously have:

1.2.7. Theorem. *(a)* $\{a\} = \{b\}$ *if and only if* $a = b$
(b) $x \in \{a\}$ *if and only if* $x = a$.

A set of the form $\{a, b\}$ where a and b are objects is called an *unordered pair*. The unordered pair $\{a, b\}$ may contain just one or two elements according to whether $a = b$ or $a \neq b$. For unordered pairs we have:

1.2.8. Theorem. *(a)* $\{a, b\} = \{b, a\}$
(b) $x \in \{a, b\}$ *if and only if* $x = a$ *or* $x = b$.

It therefore follows:

1.2.9. Theorem. $\{a, b\} = \{c, d\}$ *if and only if either (i)* $a = c$ *and* $b = d$ *or (ii)* $a = d$ *and* $b = c$.

Similarly the set $\{a, b, c\}$ may have one, two or three elements depending on the objects whose names are a, b and c. For example, $\{a, b, c\} = \{a, c\}$ if $a = b$.

1.3. SUBSETS

In Theorem 1.2.3, the criterion for the equality of two sets is given in

two conditions. But either of them can be used as a criterion for a new comparison of sets that leads to the following very useful relationship:

1.3.1 Definition. *Let A and B be sets. We say that B is a* subset *of A if A contains every element of B. In this case we write $A \supset B$ or $B \subset A$.*

Or equivalently:

1.3.2. Definition. *$A \supset B$ if and only if for every object x, $x \in A$ always follows from $x \in B$.*

We may illustrate $B \subset A$ by the so-called *Venn diagram* below (Fig. 1.1)

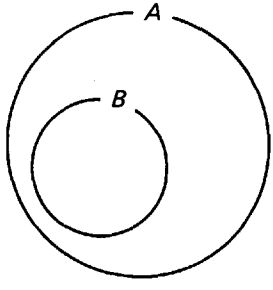


Fig. 1.1

where a smaller disc representing B is inside a bigger disc representing A. For B being a subset of A we sometimes say that B is *included in* A, B is *contained in* A as a subset, A *includes* B or A *contains* B as a subset.

An easy consequence of the above definition and the rule of extension is:

1.3.3. Theorem. *Let A and B be sets. Then $A = B$ if and only if $A \supset B$ and $B \subset A$.*

For all sets A, A itself is always a subset of A: $A \subset A$. Thus A is a trivial subset of A. Any subset B of A such that $B \neq A$ is called a *proper subset* of A. In this case we sometimes write (not too elegantly) $B \subsetneqq A$ or $A \supsetneqq B$. For example, the set of all positive even integers is a proper subset of the set of all integers.

Suppose we have a picture of an object drawn in black against a white

background. If we interchange the colour of the picture with that of the background, it will still convey to us the same object as before. (Fig. 1.2). Thus the background of a picture tells us as much as the picture does itself.

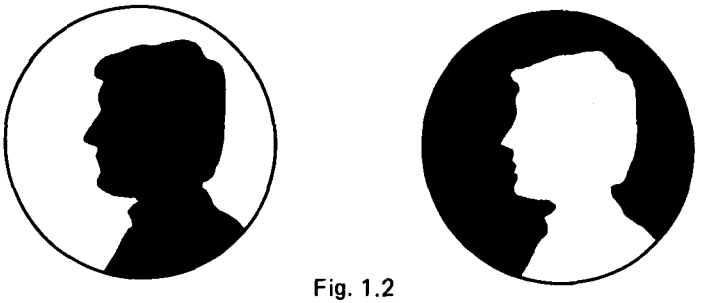


Fig. 1.2

The same holds true for sets. Here the picture of a set A is presented by all objects belonging to A (i.e. all x such that $x \in A$) whereas the background is presented by all objects not belonging to A (i.e. all y, such that $y \notin A$). In terms of background, $B \subset A$ is illustrated by Fig. 1.3,

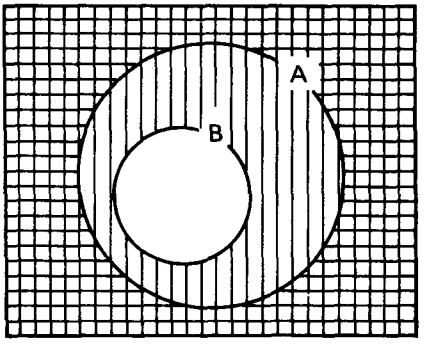


Fig. 1.3

where the horizontally shaded area is the background of A and the vertically shaded area is the background of B, the former being contained in the latter. In more precise terms we have:

1.3.4 Theorem. *Let A and B be sets. A necessary and sufficient condition*

for B to be a subset of A is that for every object y, if $y \notin A$ then $y \notin B$, or equivalently that there is no object x such that $x \in B$ and $x \notin A$.

In mathematics we often put a stroke across a symbol to denote the denial or the negation of the relationship represented by the symbol in question. For example, if a stroke is put across the symbols = and $\in$, then $\neq$ means 'not being equal to' and $\notin$ means 'not belonging to'. Similarly, $\not\subset$ means 'not being a subset of'. Given two sets A and B, a necessary and sufficient condition for $A \subset B$ is that:

(a) B contains every element of A, or
(b) every element of A belongs to B, or
(c) for every object x, if $x \in A$ then $x \in B$.

Therefore the denial (or the negation) of each of the above statements can serve as a necessary and sufficient condition for $A \not\subset B$. Let us formulate them carefully.

The negation of (a) is that:
B does not contain every element of A, or
B does not contain some elements of A, or
B does not contain at least one element of A.

The negation of (b) is that:
Not every element of A belongs to B, or
Some elements of A do not belong to B, or
At least one element of A does not belong to B.

The negation of (c) is that:
It is not true that for every object x, if $x \in A$ then $x \in B$, or
There are objects x for which it is not true that if $x \in A$ then $x \in B$, or
There is at least one object x such that $x \in A$ and $x \notin B$.

Therefore among others we have:

1.3.5. Theorem. *Let A and B be sets. Then $A \not\subset B$ if and only if there is at least one object x such that $x \in A$ and $x \notin B$.*

It is clear that given any two sets A and B, either $A \subset B$ or $A \not\subset B$.

Either of the two Venn diagrams below (Figs. 1.4 & 1.5) illustrates $A \not\subset B$.

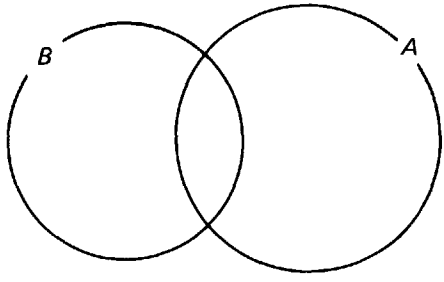


Fig. 1.4

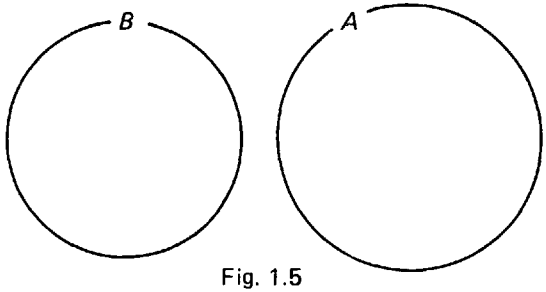


Fig. 1.5

Some formal properties of inclusion ($\subset$) are listed below.

1.3.6. Theorem. *For all sets A, B and C,*

(a) $A \subset A$ *(reflexive property of inclusion)*

(b) If $A \subset B$ *and* $B \subset C$ *then* $A \subset C$ *(transitive property of inclusion).*

We note that inclusion does not have the symmetric property, i.e. $B \subset A$ does not necessarily follow from $A \subset B$.

1.4. RULE OF SPECIFICATION

One of the most useful rules of set theory is the rule of specification.

This enables us to select elements from a given set which share a common property to form a definite subset. In order to facilitate the formulation of this rule in a useful form, we denote the proposed common property by a capital letter.

Let P be the property and x an object. Then we denote by $P(x)$ the sentence (or statement) 'x has the property P'. The truth of the statement $P(x)$ will depend on both the property P and the object x. For example, we may denote the property of being an even integer by E and the property of being a man aged 16 or over by M. Then

E (4), E (−200), M (K.T. Leung), M (P.H. Cheung)

are all true statements while

E (7), E ($-\sqrt{2}$), M (Mary Chan), M (Alice Lam)

are all false statements.

We are now in a position to formulate the rule.

1.4.1. Rule of specification. *Let A be a set and P a property that elements of A may or may not have. Then we have a subset B of A consisting of all elements of A which have the property P. This subset is denoted by*

$$B = \{x \in A : P(x)\}.$$

In other words, for any object x,

$x \in \{x \in A : P(x)\}$ if and only if $x \in A$ and $P(x)$ is a true statement.

Thus in the above notation of B, we shall still have the same subset B if we replace the letter x throughout by another letter, say y:

$$\{x \in A : P(x)\} = \{y \in A : P(y)\}.$$

Sometimes the colon (:) is replaced by a vertical stroke (|):

$$\{x \in A : P(x)\} = \{x \in A \mid P(x)\}.$$

Our new rule is nothing more than a formalization of the common procedure of selecting and grouping objects. But in this given form, it enables us to construct subsets with the precision which is required by mathematics. For example, in the set $\mathbb{Z}$ of all integers, we have the subset

$$\{x \in \mathbb{Z} : E(x)\} = \{x \in \mathbb{Z} : x = 2n \text{ for some } n \in \mathbb{Z}\}$$

of all even integers, and in the set H of all human beings, we have the subset

$$\{x \in H : M(x)\} = \{x \in H : x \text{ is a man aged 16 or over}\}$$

of all men aged 16 or over.

Let us now use the rule of specification to construct a rather extraordinary set. Consider the property P of an object being different from

itself. Thus $P(x)$ is the statement

$$x \neq x.$$

If A is any set, then

$$\{x \in A: x \neq x\}$$

is a subset of A which we shall denote by the special symbol ϕ. As a set in its own right, ϕ has the extraordinary property that it contains no element at all. To see this, we assume to the contrary that ϕ contains an object x as its element. Then this object x must satisfy both conditions:

(i) $x \in A$, and

(ii) $x \neq x$.

But (ii) is in contradiction to the reflexive law of equality (Rule 1.1.2). Therefore it cannot be true that the set ϕ contains an element x. In other words ϕ has no element at all.

If B is a set, then B contains every element of ϕ since ϕ has no element. Therefore $\phi \subset B$. Thus the set ϕ has another extraordinary property that it is a subset of every set.

It appears that if we start with another set C and follow the same procedure of construction, we might obtain another set

$$D = \{x \in C: x \neq x\}$$

which also contains no element. Similarly we might obtain yet another set F without element starting with another set E, and so forth. But this is not to be the case because all sets without any element must be identical. For if X is any set without element, then $X \supset \phi$ because X contains every element of ϕ, and $\phi \supset X$ because ϕ contains every element of X. Therefore by the rule of extension $X = \phi$. Hence there is one and only one set without element which shall be called *the empty set, the void set* or *the null set*, and the special symbol ϕ is used exclusively to denote this extraordinary set. Some properties of the empty set ϕ are:

1.4.2. Theorem. *Let A and B be sets. Then*

(a) $A = \phi$ if and only if $x \notin A$ for all objects x

(b) $B \supset \phi$ for all sets B

(c) ϕ is the only set that has no proper subset.

1.5 EXERCISE

1. Which of the following sets are equal?

$$\{a, b, c\}, \{c, b, a, c\}, \{b, c, b, a\}, \{c, a, c, b\}.$$

2. If $A = \{1, 3, 2\}$, $B = \{1, 2, 4\}$, $C = \{1, 2, 3\}$ and $D = \{1, 2\}$, then which of the following statements are false?
   (a) $A = B$;
   (b) $A = C$;
   (c) $A = D$;
   (d) $B = D$.

3. Pair the following sets so that one is the subset of the other:
   A = the set of all books,
   B = the set of all rectangles,
   C = the set of all straight lines,
   D = the set of all polygons,
   E = the set of all horizontal lines,
   F = the set of all books on geometry.

4. Let $A = \{a, b\}$. Which of the following statements are true?
   (a) $a \in A$,
   (b) $a \subset A$,
   (c) $\{a\} \in A$,
   (d) $\{a\} \subset A$.

5. Construct sets A, B, C such that $A \subset B \in C$ and $A \in B \subset C$.

6. Find four examples of a set A with the property that every element of A is a subset of A.

7. Let $A = \{a, b, c\}$. Find all the subsets of A.

8. Which of the following sets are equal?
   $$\phi, \{\phi\}, \{0\}$$

9. Let $A = \{0, \{1, 2\}\}$. Find all the subsets of A.

10. Let $A = \{\phi, \{\phi\}, \{\phi, \{\phi\}\}\}$. Which of the following statements are true?
    $\phi \subset A$, $\phi \in A$, $\{\phi\} \in A$, $\{\phi\} \subset A$, $\{\{\phi\}\} \subset A$, $\{\{\phi\}, \phi\} \subset A$, $\{\{\phi\}, \phi\} \in A$.

11. Let $\mathbb{Z}$ be the set of all integers.
If $A = \{m \in \mathbb{Z} : m = 2n \text{ for some } n \in \mathbb{Z}\}$
and $B = \{p \in \mathbb{Z} : p = 6q + 10r \text{ for some } q, r \in \mathbb{Z}\}$,
prove that $A = B$.

1.6. COMPLEMENTS

The rule of specification also permits us to remove elements from a set. More precisely, consider the two sets A and B (Fig. 1.6):

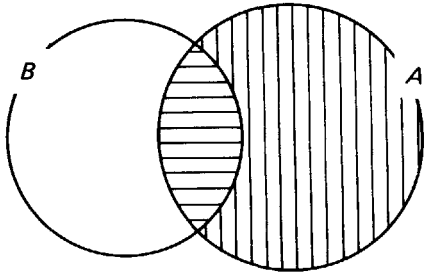


Fig. 1.6

We want to formalize the operation of removing from the set A all elements that happen to belong to the set B. Elements to be removed are objects y such that

$$y \in A \text{ and } y \in B,$$

these are in the horizontally shaded area of the above Venn diagram. Elements to be retained are objects x such that

$$x \in A \text{ and } x \notin B,$$

these are in the vertically shaded area of the Venn diagram. After separating from A the portion that is included in B we obtain the subset

$$\{x \in A : x \notin B\}$$

of A which we shall call the *complement of B in A* and denote by $A \setminus B$ (or $A - B$ as preferred by some authors). Thus we have:

1.6.1. Theorem. *Let A and B be sets. Then $x \in A \setminus B$ if and only if $x \in A$ and $x \notin B$.*

The formal properties of the operation are listed in the theorem below.

1.6.2. Theorem. *Let A and B be sets. Then*

(a) $A \setminus B \subset A$

(b) $A \setminus A = \phi$

(c) $A \setminus \phi = A$ *and* $\phi \setminus B = \phi$

(d) $A \setminus B = \phi$ *if and only if* $A \subset B$.

Proof. Statements (a), (b) and (c) are obviously true.

We exclude the trivial case where $A = \phi$.

Let A be a *non-empty* set, i.e. $A \neq \phi$. The proof of the statement (d) is divided into two parts, namely:

(i) if $A \subset B$ then $A \setminus B = \phi$, and

(ii) if $A \setminus B = \phi$ then $A \subset B$

corresponding to the 'if' and the 'only if' in the phrase 'if and only if' in the original statement (d).

Proof of (i). It follows from $A \subset B$ that if $x \in A$ then $x \in B$. Therefore there is no object x such that $x \in A$ and $x \notin B$ at the same time. Hence $A \setminus B = \{x \in A : x \notin B\} = \phi$.

Proof of (ii). It follows from $A \setminus B = \phi$ that there is no object x such that $x \in A$ and $x \notin B$. Therefore $A \subset B$ by Theorem 1.3.4. ∎

Fig. 1.7 shows that in the case where $A \subset E$, then $E \setminus A$ coincides with the background of the picture of A mentioned earlier in Section 1.3. By common sense, the background of the background of A must be the same as A; moreover, the larger the set A, the smaller the background of A.

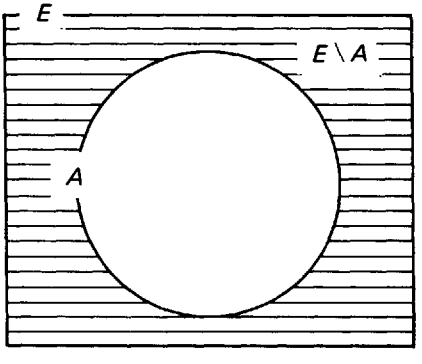


Fig. 1.7

1.6.3. Theorem. *Let A and B be subsets of E. Then*

(a) $E \setminus (E \setminus A) = A$

(b) $B \subset A$ *if and only* $E \setminus B \supset E \setminus A$.

Proof. (a) By Theorem 1.3.3, the proposed equality will be proved if we can establish the two inclusions

$$\text{(i) } A \subset E \setminus (E \setminus A) \quad \text{and} \quad \text{(ii) } A \supset E \setminus (E \setminus A).$$

(i) Let $x \in A$. Then $x \in E$ because A is a subset of E. It remains to show that $x \notin E \setminus A$. Now for the object x and the set $E \setminus A$ it is

$$\text{either } x \in E \setminus A \text{ or } x \notin E \setminus A.$$

The former case is impossible because $x \in A$ by hypothesis. Therefore $x \notin E \setminus A$. Hence $x \in E \setminus (E \setminus A)$ proving $A \subset E \setminus (E \setminus A)$.

(ii) Let $x \in E \setminus (E \setminus A)$. By definition $x \in E$ and $x \notin E \setminus A$. But for object x and the set A, it is

$$\text{either } x \notin A \text{ or } x \in A.$$

The former case is impossible, for otherwise we would have $x \in E \setminus A$ contradicting $x \notin E \setminus A$. Therefore $x \in A$. Hence $A \supset E \setminus (E \setminus A)$.

This completes the proof of (a).

(b) The proof of (b) is again divided into two parts:

(i) if $E \setminus B \supset E \setminus A$, then $B \subset A$, and

(ii) if $B \subset A$, then $E \setminus B \supset E \setminus A$

corresponding to the 'if' and the 'only if' in the phrase 'if and only if' of the original statement (b).

(i) It follows from $E \setminus B \supset E \setminus A$ that for every $x \in E$, if $x \notin A$ then $x \notin B$. On the other hand since both A and B are subsets of E, it is also true that for every $x \notin E$, if $x \notin A$ then $x \notin B$. Thus for all objects x, if $x \notin A$ then $x \notin B$. By Theorem 1.3.4 on the backgrounds of sets, we have $B \subset A$.

(ii) We may prove the inclusion using arguments similar to those used in the proof of (b)(i). Alternatively if follows from $B \subset A$ that $E \setminus (E \setminus B) \subset E \setminus (E \setminus A)$ by (a). Applying (b)(i) to $E \setminus B$ and $E \setminus A$, if follows from $E \setminus (E \setminus B) \subset E \setminus (E \setminus A)$ that $E \setminus B \supset E \setminus A$.

The proof of (b) is now complete. ∎

1.7. INTERSECTION

Two sets A and B are given as illustrated in the Venn diagram below

(Fig. 1.8), where the set A is partitioned into differently shaded areas. In the last section we cut away from A the vertically shaded area to obtain the complement $A \setminus B$ of B in A, which is represented by the horizontally shaded area in the diagram. In this section we consider the second operation of cutting away from A the horizontally shaded portion $A \setminus B$.

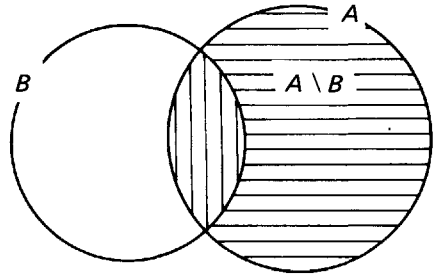


Fig. 1.8

The result of this operation is the subset

$$A \setminus (A \setminus B) = \{x \in A : x \notin A \setminus B\}$$

of A which is represented by the vertically shaded area of the diagram. Let us now study the elements of this set. An object x belongs to $\{x \in A : x \notin A \setminus B\}$ if and only if

$$x \in A \text{ and } x \notin A \setminus B.$$

On the other hand for the object x and the set B it is

$$\text{either } x \notin B \text{ or } x \in B.$$

Under the hypothesis that $x \in A$,

$$x \notin B \text{ if and only if } x \in A \setminus B, \text{ and}$$
$$x \in B \text{ if and only if } x \notin A \setminus B.$$

Therefore an object x belongs to $\{x \in A : x \notin A \setminus B\}$ if and only if

$$x \in A \text{ and } x \in B.$$

In other words,

$$\{x \in A : x \notin A \setminus B\} = \{x \in A : x \in B\}.$$

We call this set the *intersection* of the set A and the set B and denote it by $A \cap B$. Thus the intersection

$$A \cap B = \{x \in A : x \in B\}$$

is the set of all common elements of A and B.

1.7.1 Definition. *For any object x, $x \in A \cap B$ if and only if $x \in A$ and $x \in B$.*

Thus the membership of $A \cap B$ is the overlapping membership of A and B. Obviously $B \cap A = \{x \in B : x \in A\}$ is identical with $A \cap B$ above. This and other formal properties of the intersection are listed in the theorem below.

1.7.2 Theorem. *Let A, B and C be sets. Then*

(a) $A \cap A = A$

(b) $A \cap \phi = \phi$

(c) $A \cap B = B \cap A$ *(the commutative law of intersection)*

(d) $(A \cap B) \cap C = A \cap (B \cap C)$ *(the associative law of intersection).*

The proof of these four statements is left to the reader as an exercise.

Taking the two operations $\cap$ and $\setminus$ together, we see that elements of the sets A and B are partitioned into three sets $A \cap B$, $A \setminus B$ and $B \setminus A$ as illustrated in the Venn diagram below (Fig. 1.9).

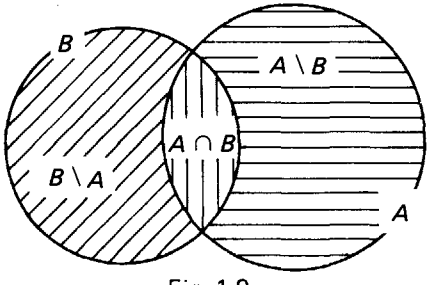
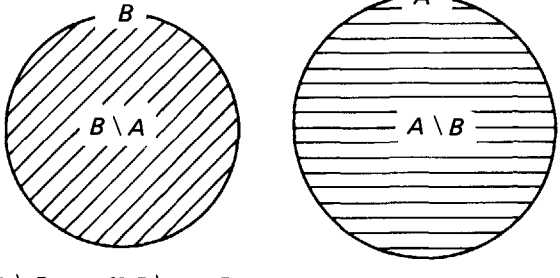


Fig. 1.9

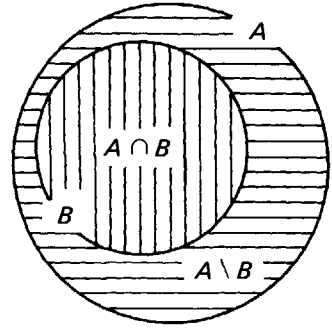
Depending on A and B, one or two of the sets $A \cap B$, $A \setminus B$ and $B \setminus A$ may be empty. Some of the possible cases are illustrated in the Venn diagrams below where the differently shaded areas have the same meaning as before (Figs. 1.10–1.12). The reader is invited to prove the statements accompanying the diagrams where the new word 'iff' is just an abbreviation of the phrase 'if and only if'.



$A \cap B = \phi$ iff $A \setminus B = A$ iff $B \setminus A = B$

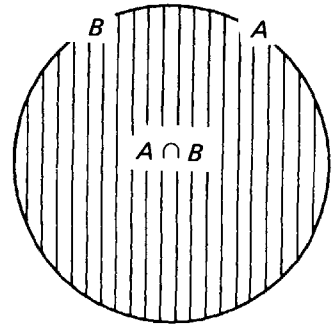
Fig. 1.10

We say that A and B are *disjoint* if they have empty intersection, i.e. $A \cap B = \phi$.



$A \cap B = B$ iff $A \supset B$ iff $B \setminus A = \phi$

Fig. 1.11



$A = B$ iff $A \setminus B = \phi$ and $B \setminus A = \phi$

Fig. 1.12

In general, intersection has the properties listed in the two theorems below in relation to inclusion.

1.7.3 Theorem. *Let A, B, A′ and B′ be sets. Then*

(a) $A \cap B \subset A' \cap B'$ *if* $A \subset A'$ *and* $B \subset B'$

(b) $A \cap B = A$ *if and only if* $A \subset B$.

1.7.4. Theorem. *Let A and B be sets. Then*

(a) $A \cap B \subset A$ *and* $A \cap B \subset B$

(b) If C is a set such that $C \subset A$ and $C \subset B$, then $C \subset A \cap B$.

The proofs of these theorems are left to the reader as an exercise. A careful reading of the last theorem reveals that (a) $A \cap B$ is included in both A and B, and (b) $A \cap B$ includes every common subset of A and B as a subset. In other words, $A \cap B$ is the largest set (in terms of inclusion) that is included in both A and B. This can be formulated also as the theorem below:

1.7.5. Theorem. *Let A, B and X be sets. Then $X = A \cap B$ if and only if the following conditions are satisfied:*

(a) $X \subset A$ *and* $X \subset B$

(b) If C is a set such that $C \subset A$ and $C \subset B$ then $C \subset X$.

1.8 UNION

We have found in $A \cap B$ the largest set (in terms of inclusion) that is included in both A and B, and this set is defined by

$$x \in A \cap B \text{ if and only if } x \in A \text{ and } x \in B.$$

Parallel to this we wish to find the smallest set that includes both A and B as subsets. By this we mean a set Y such that

(a) $A \subset Y$ and $B \subset Y$

(b) If D is a set such that $A \subset D$ and $B \subset D$, then $Y \subset D$.

It turns out that Y is defined by

$$x \in Y \text{ if and only if } x \in A \text{ or } x \in B.$$

(We observe that the connective *and* in the first definition is now replaced by the connective *or* in the second definition.) To see this we have to show that Y, whose membership is given by $x \in A$ or $x \in B$, satisfies the requirements (a) and (b) of being the smallest set that includes both A and B.

Proof. (a) If $x \in A$, then also $x \in A$ or $x \in B$. Therefore $x \in Y$ proving $A \subset Y$. Similarly $B \subset Y$.

(b) Let D be a set such that $A \subset D$ and $B \subset D$. If $x \in Y$, then $x \in A$ or $x \in B$ by definition of Y. In the former case, $x \in D$, since $A \subset D$. In the latter case $x \in D$ also, since $B \subset D$. Therefore $x \in D$, proving $Y \subset D$. ∎

Another reading of the definition reveals that Y is the set which consists of all elements of A and all elements of B; we may therefore call this set the *union* of the set A and the set B and denote it by $A \cup B$. Thus:

1.8.1. Definition. *$x \in A \cup B$ if and only if $x \in A$ or $x \in B$, or*

$$A \cup B = \{x: x \in A \text{ or } x \in B\}.$$

We may also say that the membership of $A \cup B$ is the combined membership of A and B. The union $A \cup B$ is represented by all shaded areas of the Venn diagram below (Fig. 1.13):

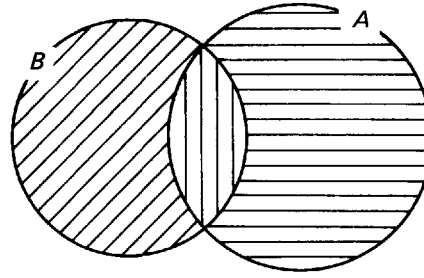


Fig. 1.13

Some formal properties of union are listed in the three theorems below whose proofs are left to the reader as an exercise.

1.8.2. Theorem. *Let A, B and C be sets. Then*

(a) $A \cup A = A$

(b) $A \cup \phi = A$

(c) $A \cup B = B \cup A$ *(the commutative law of union)*

(d) $(A \cup B) \cup C = A \cup (B \cup C)$ *(the associative law of union).*

In relation to inclusion we have:

1.8.3 Theorem. *Let A, B, A' and B' be sets. Then*

(a) $A \cup B \subset A' \cup B'$ *if* $A \subset A'$ *and* $B \subset B'$

(b) $A \cup B = A$ *if and only if* $B \subset A$

(c) $A \subset A \cup B$ *and* $B \subset A \cup B$

(d) If D is a set such that $A \subset D$ and $B \subset D$, then $A \cup B \subset D$.

In relation to intersection we have:

1.8.4. Theorem. *Let A, B and C be sets. Then the following distributive laws hold*:

(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

More intersecting and less trivial is the so-called *De Morgan's Law* which links up union, intersection and complement.

1.8.5. Theorem *(De Morgan's Law). Let A and B be subsets of a set E. Then*

(a) $E \setminus (A \cup B) = (E \setminus A) \cap (E \setminus B)$

(b) $E \setminus (A \cap B) = (E \setminus A) \cup (E \setminus B)$.

We note that these formulae resemble the usual distributive laws but with a major difference in which the symbols $\cap$ and $\cup$ interchange from one side of the equality sign to another. An illustration of De Morgan's law is given by the Venn diagram below (Fig. 1.14):

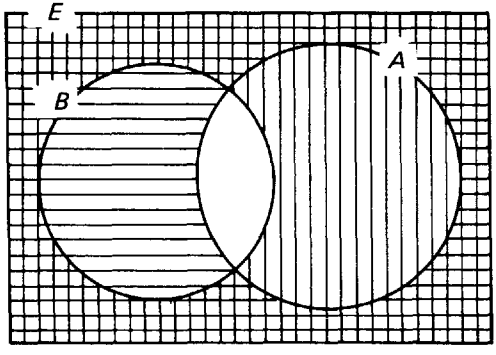


Fig. 1.14

where $E \setminus A$ is shaded horizontally and $E \setminus B$ is shaded vertically. Let us now prove the theorem.

Proof. (a) By definition

$$x \in E \setminus (A \cup B) \text{ iff } x \in E \text{ and } x \notin A \cup B.$$

By the definition of union

$$x \notin A \cup B \text{ iff } x \notin A \text{ and } x \notin B.$$

Therefore,

$x \in E \setminus (A \cup B)$ iff $x \in E$ and $x \notin A$ as well as $x \in E$ and $x \notin B$

i.e. $x \in E \setminus (A \cup B)$ iff $x \in E \setminus A$ and $x \in E \setminus B$

Hence $E \setminus (A \cup B) = (E \setminus A) \cap (E \setminus B)$.

(b) $(E \setminus A) \cup (E \setminus B) = E \setminus (E \setminus ((E \setminus A) \cup (E \setminus B)))$

$= E \setminus ((E \setminus (E \setminus A)) \cap (E \setminus (E \setminus B)))$

$= E \setminus (A \cap B)$. ■

1.9. EXERCISE

In this exercise, $A, B, C, \ldots$ denote sets.

1. Let $A = \{a, b, c, d\}$, $B = \{c, d, e\}$ and $C = \{a, e\}$. Determine
   (a) $A \cup B, B \cup C, C \cup A$ and
   (b) $A \cap B, B \cap C, C \cap A$.

2. Let $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $B = \{2, 4, 6, 8\}$, $C = \{1, 3, 5, 7, 9\}$, $D = \{3, 4, 5\}$ and $E = \{3, 5\}$. Which of these sets will be equal to the set X satisfying the following conditions?
   (a) $X \cap B = \phi$;
   (b) $X \subset D$ and $X \not\subset B$;
   (c) $X \subset A$ and $X \not\subset C$;
   (d) $X \subset C$ and $X \not\subset A$;

3. Consider the following sets of triangles:
   T = the set of all triangles,
   E = the set of all equilateral triangles,
   I = the set of all isosceles triangles,
   S = the set of all scalene triangles,
   R = the set of all right-angled traingles,
   A = the set of all acute-angled traingles.
   Which of the following statements are true?
   (a) $E \subset I \subset T$;
   (b) $S \subset E$;

(c) $A \subset R$;
(d) $I \cap R \subset A \cap E$.

4. Consider the following sets of quadrilaterals:
   Q = the set of all quadrilaterals,
   T = the set of all trapezia,
   S = the set of all squares,
   P = the set of all parallelograms,
   R = the set of all rhombuses,
   E = the set of all rectangles.
   Which of the following statements are true?
   (a) $S \subset E \subset P \subset T \subset Q$;
   (b) $R \subset P \subset T \subset Q$;
   (c) $R \cap E = S$;
   (d) $S \subset T \cap E$;
   (e) $S \cup P \subset R \cup E$.

5. Is the equality $\{a, b\} \cap \{b, c\} = \{b\}$ always true? Why?

6. If $\{A, B\} = \{C, D\}$, prove that
   (a) $A \cap B = C \cap D$,
   (b) $A \cup B = C \cup D$.

7. Give an example of sets A, B and C such that
   $$A \cap B \neq \phi,\ A \cap C \neq \phi, \text{ but } A \cap B \cap C = \phi.$$

8. If $A \cup B = A$ for any set A, prove that $B \neq \phi$.

9. If $A \cup B = A \cup C$ and $A \cap B = A \cap C$, prove that $B = C$.

10. Prove that $(A \cap B) \cup C = A \cap (B \cup C)$ if and only if $C \subset A$.

11. (a) Prove that $(A \cup B) \cap C = [A \cup (B \cap C)] \cap C$.
    (b) Deduce that $(A \cup B) \cap C = A \cup (B \cap C)$ if and only if $A \subset C$.

12. If $A \subset B$, prove that there is a unique subset S of B such that $S \cup A = B$ and $S \cap A = \phi$.

13. Are the following statements true for all sets A, B and C? Explain.

(a) If $A \notin B$ and $B \notin C$, then $A \notin C$.
(b) If $A \neq B$ and $B \neq C$, then $A \neq C$.
(c) If $A \in B$ and $B \not\subset C$, then $A \notin C$.
(d) If $A \subset B$ and $B \subset C$, then $C \not\subset A$.
(e) If $A \subset B$ and $B \in C$, then $A \notin C$.
(f) If $A \cap C \subset B$, then $(A \cap B) \cup (B \cap C) = B$.
(g) If $A \cap C \in B$, then $A \in B \cup C$.

14. Prove that the following statements are equivalent:
(a) $A \subset B$,
(b) $A \cap B = A$ and
(c) $A \cup B = B$.

15. Let S be a collection of sets such that $A \setminus B \in S$ for any $A, B \in S$. If A and $B \in S$, prove that $A \cap B \in S$.

16. Let S be a collection of sets. Define

$$S^0 = \{A \setminus B : A \text{ and } B \in S\}$$

(a) Prove $S^0 \subset (S^0)^0$.
(b) Give an example to show that it is possible to have $S^0 \neq (S^0)^0$.

17. If A is a set such that $A \notin A$, determine of which of the following sets is A an element, a subset, and neither an element nor a subset:
(a) $\{\{A\}, B\}$ $\quad (A \neq B)$,
(b) A,
(c) $\{A\} \setminus \{\{A\}\}$,
(d) $\{A\} \cup A$,
(e) $\{A\} \cup \{\phi\}$.

18. If $A \subset B$, prove that $A \cup (B \setminus A) = B$.

19. Prove that (a) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$
(b) $A \setminus B = A \setminus (A \cap B)$.

20. If A and $B \subset E$, prove that
(a) $A \setminus B = A \cap (E \setminus B)$,
(b) $A \setminus (E \setminus B) = A \cap B$,
(c) $(E \setminus A) \setminus (E \setminus B) = B \setminus A$,
(d) $A \cup [(E \setminus A) \cap B] = A \cup B$,
(e) $A \cap [(E \setminus A) \cup B] = A \cap B$.

21. If A and $B \subset E$ such that $A \cap B = \phi$, prove that $A \subset E \setminus B$.

22. Let A, B and C be subsets of E. Prove that
    (a) $E \setminus (A \cup B \cup C) = (E \setminus A) \cap (E \setminus B) \cap (E \setminus C)$ and
    (b) $E \setminus (A \cap B \cap C) = (E \setminus A) \cup (E \setminus B) \cup (E \setminus C)$.

23. Prove that (a) $(A \setminus B) \setminus C = (A \setminus C) \setminus (B \setminus C)$,
    (b) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$,
    (c) $A \setminus (B \cup C) = (A \setminus B) \setminus C$,
    (d) $(A \setminus B) \cup (A \cap C) = A \setminus (B \setminus C)$,
    (e) $(A \setminus B) \cup (A \setminus C) = A \setminus (B \cap C)$.

24. Define $A \mid B = E \setminus (A \cap B)$, where A and B are subsets of E.
    (a) Prove that (i) $E \setminus A = A \mid A$
    and (ii) $A \cap B = (A \mid B) \mid (A \mid B)$.
    (b) Express $\cup$ (union) in terms of $\mid$ alone

*25. Let $A_1, A_2, \ldots, A_n$ be n given sets.
    (a) For $i = 1, 2, \ldots, n$, find a subset S_i of A_i such that $S_i \subset S_j$ for $i > j$ and $A_1 \cap A_2 \cap \ldots \cap A_i = S_1 \cap S_2 \cap \ldots \cap S_i$.
    (b) For $i = 1, 2, \ldots, n$, find a subset T_i of A_i such that $T_i \cap T_j = \phi$ if $i \neq j$ and $A_1 \cup A_2 \cup \ldots \cup A_i = T_1 \cup T_2 \cup \ldots \cup T_i$.

1.10. ORDERED PAIRS AND CARTESIAN PRODUCT

In terms of number of elements, the empty set ϕ is the smallest set possible and there is just one such set with no element. The next larger sets are the singletons; these are sets with exactly one element. They are written in the form

$$\{a\}$$

where a is an object and the only element of the set $\{a\}$. The unordered pairs

$$\{a, b\}$$

where $a \neq b$ are the next larger sets with exactly two elements. Properties of these small sets are formulated in Theorems 1.2.7, 1.2.8 and 1.2.9. In particular we note that in the unordered pair $\{a, b\}$, the elements a and b play an identical role as members of the unordered pair $\{a, b\}$.

We are now interested in constructing a set from a pair of objects a and b in which a and b play rather different roles. Moreover we would like this set to have the smallest possible number of elements. Obviously this number cannot be zero nor one; therefore it has to be at least two.

Starting with objects a and b, and using all available techniques such as $\{\ \}$, $\cap$, $\cup$ and $\setminus$ once, we can only obtain three non-empty sets,

$$\{a\},\ \{b\},\ \{a, b\}.$$

But none of these first-generation species satisfies our requirement. On the other hand, we can use these sets as well as the original a and b as objects to construct many more sets using the same techniques. Among them we have, for example, the following singletons

$$\{\{a\}\},\ \{\{b\}\},\ \{\{a, b\}\}, \ldots$$

and the following unordered pairs,

$$\{a, \{a\}\},\ \{a, \{b\}\},\ \{\{a\},\ \{b\}\},\ \{\{a\},\ \{a, b\}\}$$
$$\{\{a, b\},\ \{b\}\}, \ldots$$

Among these second-generation species, we do find sets in which a and b have different roles to play. Take, for example, the fourth one on the second list,

$$\{\{a\},\ \{a, b\}\}$$

where a appears twice and b only once. Thus, at least on appearance, a and b have asymmetric roles. This is also easily confirmed by an application of Theorem 1.2.9. that

$$\text{if } a \neq b \text{ then } \{\{a\},\ \{a, b\}\} \neq \{\{b\},\ \{b, a\}\}.$$

Most importantly we can prove the following crucial theorem.

1.10.1. Theorem. *Let a, b, c and d be objects. Then*

$$\{\{a\},\ \{a, b\}\} = \{\{c\},\ \{c, d\}\} \quad \textit{if and only if } a = c \text{ and } b = d.$$

Proof. By Theorem 1.2.9, if $a = c$ and $b = d$, then $\{\{a\},\ \{a, b\}\} = \{\{c\},\ \{c, d\}\}$. Thus the 'if' part of the theorem is trivial. Let us prove the 'only if' part. Assume that $\{\{a\},\ \{a, b\}\} = \{\{c\},\ \{c, d\}\}$. By Theorem 1.2.9 again, there are the following two possible cases for the elements of these sets:

(i) $\{a\} = \{c\}$ and $\{a, b\} = \{c, d\}$, or

(ii) $\{a\} = \{c, d\}$ and $\{a, b\} = \{c\}$.

In case (i), it follows from $\{a\} = \{c\}$ that $a = c$. $a = c$ together with $\{a, b\} = \{c, d\}$ must imply that $b = d$. Therefore in this case we have $a = c$ and $b = d$.

In case (ii), it follows from $\{a\} = \{c, d\}$ that $a = c = d$ and it follows from $\{a, b\} = \{c\}$ that $a = b = c$. Thus $a = b = c = d$. Hence, also $a = c$ and $b = d$. The proof of the theorem is now complete. ■

We have therefore succeeded in our construction of a set with the least number of elements in which a and b have asymmetric roles. Furthermore, the last theorem shows that the set $\{\{a\}, \{a, b\}\}$ behaves exactly like a pair of coordinates of a point in plane analytic geometry. For this reason, we shall adopt the well-known notation of analytic geometry and write

$$(a, b) = \{\{a\}, \{a, b\}\}$$

and call this set the *ordered pair* of a and b.

We can now rewrite the last theorem in this notation:

1.10.2. Theorem. *Let a, b, c and d be objects. Then $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.*

Finally we observe that $\{\{a\}, \{a, b\}\}$ is not the only one among the second-generation species that can satisfy our requirement. For example, $\{\{b\}, \{a, b\}\}$ also satisfies our requirement.

If we collect all ordered pairs (a, b) where the first coordinate a is taken from a set A and the second coordinate b is taken from a set B, then we have a set

$$\{(a, b) : a \in A \text{ and } b \in B\}$$

of all such ordered pairs. This set is clearly the analogue of the coordinate plane of analytic geometry. We shall call this set the *Cartesian product* of the set A and the set B after the French philosopher and mathematician René Descartes (1596-1650) and denote it by $A \times B$. Thus:

1.10.3. Definition. $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.

For example,

$$\{a\} \times \{b\} = \{(a, b)\}$$
$$\{a\} \times \{b, c\} = \{(a, b), (a, c)\}$$
$$\{x, y\} \times \{s, t\} = \{(x, s), (x, t), (y, s), (y, t)\}$$
$$A \times \phi = \phi = \phi \times B.$$

In general the Cartesian product of sets is not commutative:

$$A \times B \neq B \times A \text{ unless } A = B \text{ or } A = \phi \text{ or } B = \phi.$$

In analytic geometry of space we use three coordinates for each point. Analogously, if a, b and c are objects we define an *ordered triple* (a, b, c) in terms of ordered pairs by putting

$$(a, b, c) = ((a, b), c)$$

i.e. it is an ordered pair whose first coordinate is an ordered pair. Clearly we can prove that for objects, a, b, c, a', b', and c',

$$(a, b, c) = (a', b', c') \text{ iff } a = a', b = b' \text{ and } c = c'.$$

The Cartesian product $A \times B \times C$ of three sets A, B and C is therefore

$$A \times B \times C = (A \times B) \times C = \{(a, b, c) : a \in A, b \in B, c \in C\}$$

Step by step we can define *ordered n-tuples*

$$(a_1, a_2, \ldots, a_n)$$

in a similar manner. For ordered n-tuples we have

$$(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n) \text{ iff } a_i = b_i \text{ for all } i = 1, 2, \ldots, n.$$

The set of all ordered n-tuples $(a_1, a_2, \ldots, a_n)$ where the *i-th coordinate* a_i is taken from a set A_i for $i = 1, 2, \ldots, n$ is then the Cartesian product

$$A_1 \times A_2 \times \ldots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_i \in A_i; i = 1, 2, \ldots, n\}$$

1.11. EXERCISE

In this exercise, A, B, C, D denote sets.

1. Let $A = \{1, 2, 3\}$ and $B = \{a, b\}$. Find $A \times B$ and $B \times A$.

2. Use the definition $(a, b) = \{\{a\}, \{a, b\}\}$ to prove that, for any object a,
$$\{a\} \times \{a\} = \{\{\{a\}\}\}.$$

3. An ordered pair is by definition a set. Show by an example that not every ordered pair has two elements.

4. Give an example of sets A, B, C, D with
$$(A \cup B) \times (C \cup D) \neq (A \times C) \cup (B \times D).$$

5. Prove that $(A \cup B) \times C = (A \times C) \cup (B \times C)$.

6. If $A \times A = B \times B$, prove that $A = B$.

7. If $A \cap B = \phi$, prove that $(A \times B) \cap (B \times A) = \phi$.

8. If $A \neq \phi$ and $A \times B = A \times C$, prove that $B = C$.

9. If $A \neq \phi$ and $B \neq \phi$, prove that $(A \times B) \cup (B \times A) = C \times C$ if and only if $A = B = C$.

10. (a) Prove that $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.
    (b) Deduce from (a) that
    $$(A \cap B) \times C = (A \times C) \cap (B \times C)$$
    and
    $$A \times (B \cap C) = (A \times B) \cap (A \times C).$$

1.12. ONE-TO-ONE CORRESPONDENCE

In mathematics we are used to comparing things. Given two objects a and b, we can only compare them for the purpose of finding out if they are equal ($a = b$) or they are distinct ($a \neq b$). Given two sets A and B, in addition to the comparison for identity, we can also compare one with the other in order to find out if $A \subset B$ or $A \not\subset B$. The criteria for both comparisons are given in terms of membership, namely,

for $A = B$: $x \in A$ if and only if $x \in B$

for $A \subset B$: if $x \in A$ then $x \in B$.

If A and B are both finite sets with not too many elements, a procedure of the comparison for inclusion may be carried out as follows. As a first step we pick one element, say x, of A and search for an identical copy of x among the elements of B. If this cannot be found then we conclude that $A \not\subset B$ and the procedure stops. On the other hand, if a copy of x is found in B, then we proceed to the second step by picking another element, say y ($x \neq y$), of A and searching for an identical copy of y among the elements of B. If this cannot be found, then we conclude that $A \not\subset B$ and we stop; otherwise we proceed to the third step, and so forth. Thus the procedure comes to a halt whenever no identical copy of an element of A can be found in B, or when we have found identical copies for all elements of A. In the former case, we conclude that $A \not\subset B$; in the latter case, $A \subset B$. If it so happens that the last element of A is matched by its identical copy, which is the last element of B, then our conclusion will be $A = B$.

We now modify slightly the procedure above to make another kind of comparison. The modification is, that instead of searching for an identical

copy of x in B for each $x \in A$, which is very restrictive and difficult at times, we simply match each element of A by some element of B. We begin with picking an element x of A and matching it by any element of B, which once selected will be denoted by $f(x)$ to indicate that x and $f(x)$ are now matched. If there is no more element in A or in B, we stop; otherwise we proceed to the next step. The second step is to pick another element y $(x \neq y)$ of A and match it by any unused element $f(y)$ (i.e. $f(x) \neq f(y)$) in B, and so on. We continue with this procedure as long as there are still unused elements left in both A and B. When the procedure comes to a halt, then exactly one of the following three alternatives is the ultimate outcome of the comparison:

(i) elements of A are used up before those of B,

(ii) elements of B are used up before those of A, or

(iii) elements of A and B are used up at the same time.

Clearly if the outcome is (iii) we would say that the sets A and B are fully matched by pairing each $x \in A$ with $f(x) \in B$, or we may say that A and B are in a one-to-one correspondence.

We now try to put this notion of one-to-one correspondence in precise terms. At first we take note that by matching an element $x \in A$ by an element $f(x) \in B$, we get an element $(x, f(x))$ of the Cartesian product $A \times B$. If we let x run through all elements of A, the ordered pairs $(x, f(x))$ will form a subset of $A \times B$. We now lay down the definition of a correspondence.

1.12.1. Definition. *Let A and B be sets. Any subset φ of $A \times B$ is called a* correspondence *from A to B. We say that $x \in A$ and $y \in B$ correspond to each other if $(x, y) \in \varphi$.*

1.12.2. *Example.* The set $\{(x, f(x)) \in A \times B : x \in A\}$ of all matched pairs of the last procedure is a correspondence from A to B.

1.12.3. *Example.* Let S be the set of all points in the plane and T the set of all lines. The relation of incidence (the point P is incident with the line g if P lies on g) defines a correspondence

$$\{(P, g) \in S \times T : P \in g\}$$

from points to lines. To each point P correspond all lines through P and to each line g correspond all points on g. This is an example of a 'many-to-

many' correspondence where to each element of S correspond many elements of T and vice versa.

1.12.4. *Example.* If b_0 is a fixed element of B, then

$$\{(x, b_0) \in A \times B : x \in A\}$$

is an example of a 'many-to-one' correspondence where all elements of A correspond to one and the same element b_0 of B.

An important special case is the one-to-one correspondence.

1.12.5. Definition. *A correspondence φ from a set A to a set B is called a* one-to-one correspondence *if there corresponds to each $x \in A$* just one *$y \in B$ and to each $y \in B$* just one *$x \in A$, i.e.*

(i) to each $x \in A$ there is one $y \in B$ such that $(x, y) \in \varphi$
(ii) if (x, y) and (x, y') both belong to φ, then $y = y'$
(iii) to each $y \in B$ there is one $x \in A$ such that $(x, y) \in \varphi$
(iv) if (x, y) and (x', y) both belong to φ, then $x = x'$.

A one-to-one correspondence is also called a *one-one* or a *1-1* correspondence or a *bijection.* If $\varphi \subset A \times B$ is a one-to-one correspondence then the unique element $y \in B$ that corresponds to an element $x \in A$ is also denoted by $\varphi(x)$ and we may also denote the correspondence φ by $x \to \varphi(x)$.

1.12.6. *Example* The correspondence

$$\{(x, f(x)) \in A \times B : x \in A\}$$

defined by the pairing of our last procedure is a one-to-one correspondence if and only if the outcome of the comparison is the alternative (iii).

1.12.7. *Examples.* A bijection between the set of all letters of the English alphabet and the set of the first 26 positive integers is given by the correspondence

a	b	c	d	.	.	.	x	y	z
↓	↓	↓	↓				↓	↓	↓
1	2	3	4	.	.	.	24	25	26

There are, of course, many other bijections between these two sets, e.g. the correspondence

$$\begin{array}{ccccccccccc} a & b & c & d & . & . & . & x & y & z \\ \downarrow & \downarrow & \downarrow & \downarrow & & & & \downarrow & \downarrow & \downarrow \\ 3 & 4 & 5 & 6 & . & . & & 26 & 1 & 2 \end{array}$$

1.12.8. *Examples.* The following correspondence from the set of all positive even integers to the set of all positive integers

$$\begin{array}{ccccccccccccccc} & 2 & & 4 & & 6 & . & . & . & . & . & . & 2n & . & . & . & . & . & . \\ & \downarrow & & \downarrow & & \downarrow & & & & & & & \downarrow & & & & & & \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & . & . & . & 2n-1 & & 2n & & 2n+1 & & . & . & . \end{array}$$

fails to be a bijection; while the correspondence

$$\begin{array}{ccccccccccccc} 2 & 4 & 6 & 8 & . & . & . & 2n & . & . & . & . \\ \downarrow & \downarrow & \downarrow & \downarrow & & & & \downarrow & & & & \\ 1 & 2 & 3 & 4 & . & . & . & n & . & . & . & . \end{array}$$

is a bijection.

1.12.9. *Examples.* Let $\mathbb{Z}$ be the set of all integers. Then

$$\{(x, -x) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$
$$\{(x, x - 100) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$

are 1-1 correspondences from $\mathbb{Z}$ to $\mathbb{Z}$, while

$$\{(x, x^2) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$

and

$$\{(x, x^3 + 2x^2 - 8) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$

are not.

1.13. MAPPINGS

Many correspondences that fail to be a one-to-one correspondence may, nevertheless, be very useful correspondences. For example, the first correspondence of 1.12.8

$$\begin{array}{ccccccccccccccc} & 2 & & 4 & & 6 & . & . & . & . & . & . & 2n & . & . & . & . & . & . \\ & \downarrow & & \downarrow & & \downarrow & & & & & & & \downarrow & & & & & & \\ 1 & 2 & 3 & 4 & 5 & 6 & . & . & . & 2n-1 & & & 2n & & 2n+1 & & . & . & . \end{array}$$

certainly helps us to identify the set of all positive even integers as a subset of the set of all positive integers, and the correspondences

$$\{(x, x^2) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$
$$\{(x, x^3 + 2x^2 - 8) \in \mathbb{Z} \times \mathbb{Z} : x \in \mathbb{Z}\}$$

of 1.12.9 are just the polynomials

$$x^2 \text{ and } x^3 + 2x^2 - 8$$

in disguise. On closer examination, we find that all three correspondences above satisfy conditions (i) and (ii) but not both (iii) and (iv) of Definition 1.12.5 of a one-to-one correspondence. This leads us to formulate the next definition of the notion of mapping which is fundamental in any course of higher mathematics.

1.13.1. Definition. *A mapping or a function from a set A to a set B is a correspondence $\varphi \subset A \times B$ that satisfies the two conditions below:*

(i) to each $x \in A$ there is one $y \in B$ such that $(x, y) \in \varphi$

(ii) if (x, y) and (x, y') both belong to φ, then $y = y'$.

In other words, a mapping is a correspondence such that to each $x \in A$ there corresponds exactly one $y \in B$. If φ is a mapping, we write $\varphi: A \longrightarrow B$ or $A \xrightarrow{\varphi} B$ and call A the *domain* and B the *range* of φ. The unique element $y \in B$ that corresponds to $x \in A$ is called the *image* of x and is written $\varphi(x)$. We also write $x \longrightarrow y$ or $x \longrightarrow \varphi(x)$ to indicate the correspondence between x and its image $y = \varphi(x)$. Often the subset of B that consists of all the images of elements of A, i.e. the subset

$$\{y \in B : y = \varphi(x) \text{ for some } x \in A\} = \{\varphi(x) : x \in A\}$$

of B is also called the *total image* of the mapping φ (or the *direct image* of A under φ) and is denoted by $\operatorname{Im} \varphi$ or $\varphi[A]$.

Obviously all three correspondences above are mappings, and clearly a bijection is a particular type of mappings. There are mappings $\varphi: A \longrightarrow B$ that further satisfy either or both of the conditions

(iii) to each $y \in B$ there is one $x \in A$ such that $\varphi(x) = y$, and

(iv) if $\varphi(x) = \varphi(x')$ then $x = x'$

of Definition 1.12.5. It is useful to consider these two conditions separately.

1.13.2. Definition. *A mapping $\varphi: A \longrightarrow B$ is called* surjective, *or a mapping* onto *B if every element of B is an image, i.e.* $\operatorname{Im} \varphi = B$.

1.13.3. Definition. *A mapping $\varphi: A \longrightarrow B$ is called* injective *or* one-one *if distinct elements of A have distinct images, i.e. $x \neq x'$ implies $\varphi(x) \neq \varphi(x')$.*

Thus a mapping is *bijective* or is a bijection if it is both surjective and

injective. The *inclusion mapping* $2n \longrightarrow 2n$ of the set of all positive even integers into the set of all positive integers of Example 1.12.8 is injective but not surjective. The *polynomial functions* $x \to x^2$ and $x \to x^3 + 2x - 8$ are neither injective nor surjective. Examples of mappings that are surjective but not necessarily injective are the *projections* of the Cartesian product $pr_1 : A \times B \longrightarrow A$ and $pr_2 : A \times B \longrightarrow B$ defined by the correspondences $(a, b) \longrightarrow a$ and $(a, b) \longrightarrow b$ respectively.

1.14. EXERCISE

In this exercise, $\mathbb{R}$ denotes the set of all real numbers, $\mathbb{Z}$ denotes the set of all integers and $\mathbb{N}$ denotes the set of all natural numbers $= \{0, 1, 2, 3, \ldots\}$.

1. Are the following correspondences mappings? Explain.
   (a) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x + y = 1\}$;
   (b) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : xy = 2\}$;
   (c) $\{(m, n) \in \mathbb{Z} \times \mathbb{Z} : m - n = 3N \text{ for some } N \in \mathbb{Z}\}$.

2. Determine whether the following correspondences on $\mathbb{R}$ are mappings:
   (a) $\{((x, y), f[(x, y)]) \in \mathbb{R}^2 \times \mathbb{R}^2 : f[(x, y)] = (x + 1, y + 2)\}$;
   (b) $\{((x, y), g[(x, y)]) \in \mathbb{R}^2 \times \mathbb{R}^2 : g[(x, y)] = (x \cos\theta - y \sin\theta,$ $x \sin\theta + y \cos\theta)$, where θ is a fixed real number.$\}$

3. Show, by means of an example, that for a mapping f, the sets $f[A \cap B]$ and $f[A] \cap f[B]$ may be distinct.

4. If there exists a mapping from a set A to a set B, which is not injective, prove that $A \neq \phi$ and $B \neq \phi$.

5. If there exists a mapping from a set A to a set B, which is not surjective, prove that $B \neq \phi$.

6. The mapping $f : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ maps an ordered pair of natural numbers onto a single natural number by the rule $f[(m, n)] = 2^{m+1}(2n + 1)$. Show that f is an injective mapping.

7. Let E be the set of all even natural numbers. For each $n \in \mathbb{N}$, let $f(n) = 2n$. Prove that the mapping $f : \mathbb{N} \longrightarrow E$ is injective.

8. Prove that the mapping $f : \mathbb{N} \longrightarrow \mathbb{N}$ defined by
$$f(n) = n - 1 \text{ for } n \geqslant 1$$
$$\text{and } f(0) = 0$$
is not injective.

9. Prove that the mapping $f : \mathbb{N} \longrightarrow \mathbb{N}$ defined by
$$f(n) = n + 2$$
is not surjective.

10. Define a mapping $f : \mathbb{N} \longrightarrow \mathbb{N}$ by
$$f(n) = \text{sum of the digits of } n \text{ for any } n \in \mathbb{N}.$$
Is f surjective? Is f injective? Explain.

11. Prove that the mapping $f : \mathbb{R} \longrightarrow [-1, 1]$ defined by
$$f(x) = \cos x$$
is surjective but not injective.

12. Prove that the mapping $f : \mathbb{R} \longrightarrow \{-\pi, 0, \pi\}$ defined by
$$f(x) = \begin{cases} \pi & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\pi & \text{if } x < 0 \end{cases}$$
is surjective but not injective.

13. Let $A = \{1, 2, 3, \ldots, n\}$. Prove that
(a) if a mapping $f : A \longrightarrow A$ is surjective, then it is injective,
(b) if a mapping $g : A \longrightarrow A$ is injective, then it is surjective.

14. Prove that the mapping $f : \mathbb{N} \longrightarrow \mathbb{Z}$ defined by
$$f(n) = \begin{cases} -\dfrac{n}{2} & \text{if } n \text{ is even} \\ \dfrac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$
is bijective.

15. Let N be a fixed positive integer. Prove that the mapping $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ given by
$$f(n) = \begin{cases} n + N & \text{if } n \text{ is divisible by } N \\ n & \text{if } n \text{ is not divisible by } N \end{cases}$$
is bijective.

16. Give an example of non-empty sets A, B and C with the property that there are injective mappings

$$A \longrightarrow B \longrightarrow C \longrightarrow A$$

none of which are bijective.

17. Let $f : A \longrightarrow A$ and $I_A : A \longrightarrow A$ such that $I_A(a) = a$ for each $a \in A$ (I_A is called the identity mapping of A).
    (a) If $f \subset I_A$, prove that $f = I_A$.
    (b) If $I_A \subset f$, prove that $f = I_A$.

*18. If $f : A \longrightarrow B$ and X, Y are subsets of A, prove that
    (a) $f[X \cap Y] \subset f[X] \cap f[Y]$;
    (b) $f[X \cup Y] = f[X] \cup f[Y]$;
    (c) f is injective if and only if
        $f[X \cap Y] = f[X] \cap f[Y]$ for any $X, Y \subset A$;
    (d) f is surjective if and only if
        $B \setminus f[X] \subset f[A \setminus X]$ for any $X \subset A$;
    (e) f is bijective if and only if
        $f[A \setminus X] = B \setminus f[X]$ for any $X \subset A$.

*19. Let X and Y be two non-empty sets. The *disjoint union* of them is defined by

$$X + Y = (X \times \{0\}) \cup (Y \times \{1\}).$$

Define $f : X \longrightarrow X + Y$ and $g : Y \longrightarrow X + Y$ by

$$f(x) = (x, 0) \text{ for any } x \in X \qquad \text{and}$$
$$g(y) = (y, 1) \text{ for any } y \in Y.$$

    (a) Show that $f[x] \cap g[Y] = \phi$.
    (b) Show that f and g are injective.
    (c) If $h : X + Y \longrightarrow X \cup Y$ is defined by $h[(a, b)] = a$, prove that
        (i) h is surjective, and
        (ii) h is injective if and only if $A \cap B = \phi$.

*20. If X and Y are sets, denote by Y^X the set of all mappings from X to Y.

(a) Prove that, for any sets A, B and C, it is possible to find a bijection between

(i) $(A \times B)^C$ and $A^C \times B^C$;

(ii) $A^{B \cup C}$ and $A^B \times A^C$ if $B \cap C = \phi$.

(b) If there is a bijection between $A^{B \cup C}$ and $A^B \times A^C$, is it necessary for $B \cap C$ to be empty? Why?

2. *Mathematical Induction*

In this chapter we shall study in some detail a very important and useful property of natural numbers — *the principle of mathematical induction*, upon which is based the well-known technique of the proof by induction. Instead of just recapitulating the standard procedure of this particular method of proof which you have probably learnt and mastered, we shall go deeper into the subject and try to understand the nature of mathematical induction. Indeed we shall see how this principle is derived from a very simple property of natural numbers — *the well-ordering property.*

2.1. A PROOF BY INDUCTION

We are all familiar with the following formula for the sum of an arithmetic progression:

$$1 + 2 + \ldots + n = \frac{n}{2}(n + 1).$$

This formula can be proved in a most straight-forward manner as follows:

First proof. Let n be an arbitrary natural number. Then

$$1 + 2 + \ldots + n = \frac{1}{2}[(1 + 2 + \ldots + n) + (1 + 2 + \ldots + n)]$$

$$= \frac{1}{2}[(n + (n - 1) + \ldots + 2 + 1) + (1 + 2 + \ldots + (n - 1) + n)]$$

$$= \frac{1}{2}[(n + 1) + (n + 1) + \ldots + (n + 1) + (n + 1)] \qquad (n \text{ summands})$$

$$= \frac{n}{2}(n + 1).$$ ■

Alternatively we can prove the same formula by induction as follows:

Second proof. Usually this proof is in three parts.

(a) For $n = 1$, we verify the formula by substitution. Thus $1 = \frac{1}{2}(1 + 1)$.

(b) Suppose that the formula is true for some natural number $n \geqslant 1$. We proceed to show that it is also true for the next natural number $n + 1$. In other words, assuming

$$1 + 2 + \ldots + n = \frac{n}{2}(n + 1)$$

for a particular natural number n, we have to prove that

$$1 + 2 + \ldots + n + (n + 1) = \frac{(n + 1)}{2}[(n + 1) + 1].$$

Now $1 + 2 + \ldots + n + (n + 1) = [1 + 2 + \ldots + n] + (n + 1)$

$$= \frac{n}{2}(n + 1) + (n + 1) = (n + 1)(\frac{n}{2} + 1)$$

$$= \frac{n + 1}{2}(n + 2) = \frac{(n + 1)}{2}[(n + 1) + 1].$$

(c) Since the formula holds for $n = 1$ and, on the assumption of its validity for some particular n, also for the next number $n + 1$, we may conclude, by virtue of the principle of mathematical induction, that the formula holds for every natural number n. ■

Either proof establishes the validity of the formula for the sum of an arithmetic progression. In the first proof, n is an arbitary natural number, and the formula is derived from the familiar laws of arithmetic, such as the commutative, associative and distributive laws. Therefore we call this a *deductive proof*. It is said that as a primary pupil, C.F. Gauss used the same idea as that of this proof to find $1 + 2 + \ldots + 100 = 5{,}050$ when he was told to do the sum in class.

Let us analyse the individual parts of the second proof. Part (a) states the validity of the most trivial case when $n = 1$ and nothing further. Part (b) is more general, but by itself it does not tell us $1 + 2 + \ldots + 100 = 5{,}050$, until we know for sure that $1 + 2 + \ldots + 99 = 4{,}950$. Thus these two parts when taken separately do not say much on the formula that we want to prove. When taken together they do tell us a great deal. For example, knowing that the formula is true for $n = 1$ by (a), we may conclude that it holds for $n = 2$ by (b). Applying (b) to the already established case for $n = 2$, we know that the formula is also true for $n = 3$. Thus using such argument ninety-nine times, we get $1 + 2 + \ldots + 100 = 5{,}050$. Similarly, using it nine hundred and ninety-nine times, we obtain $1 + 2 + \ldots + 1{,}000 = 500{,}500$, and so on. Therefore these two parts

provide us with the necessary mechanism to verify the formula for each individual case.

If there were only a finite number of such individual cases (for instance, if we were only asked to prove the formula for the first ten thousand natural numbers n), then the two parts of the proof would be sufficient, for we could simply go through the procedures mechanically, even though it might take a very long time. However, for our present problem, we have an infinite number of individual cases (one for each natural number $n = 1, 2, 3, \ldots$) and therefore parts (a) and (b) alone are not sufficient to establish the validity of the formula for all n. It is precisely for this reason that part (c) is needed to complete the proof. Part (c) appeals directly to the principle of mathematical induction, which has the effect that on the basis of parts (a) and (b), our formula will hold by a single stroke for each of the infinitely many individual cases.

The pattern of this second proof is that we conclude that the formula is generally true on the basis of the validity of the formula in individual cases. It is therefore called a *proof by induction.*

The principle of mathematical induction is a very powerful and indispensable tool of mathematics. We shall see in the subsequent sections how it is properly used in a variety of problems and how it can be derived from an apparently simple property of natural numbers.

2.2. THE WELL–ORDERING PRINCIPLE

By natural numbers we mean non-negative integers, i.e. the whole numbers 0, 1, 2, 3, ... It is customary to denote the entire collection of natural numbers by $\mathbb{N}$ and we call it *the set of natural numbers.* We compare natural numbers by their magnitudes and know that given any two natural numbers a and b, either $a < b$ or $a = b$ or $a > b$.

Besides the entire set $\mathbb{N}$ of natural numbers we shall also consider subsets of $\mathbb{N}$. For example we have:

S_1, the set of all even numbers;

S_2, the set of all integers greater than 100;

S_3, the set of all whole numbers between and including 6 and 71;

S_4, the set of all natural numbers less than or equal to 120;

S_5, the set of natural numbers greater than 71 and less than 65;

and so on. In the notation of set theory we may write these subsets as

$S_1 = \{x \in \mathbb{N} \mid x \text{ is even}\} = \{0, 2, 4, 6, \ldots\}$
$S_2 = \{x \in \mathbb{N} \mid x > 100\} = \{101, 102, 103, \ldots\}$
$S_3 = \{x \in \mathbb{N} \mid x \geqslant 6 \text{ and } x \leqslant 71\} = \{x \in \mathbb{N} \mid 6 \leqslant x \leqslant 71\}$
$= \{6, 7, 8, \ldots, 70, 71\}$
$S_4 = \{x \in \mathbb{N} \mid x \leqslant 120\} = \{0, 1, 2, \ldots, 119, 120\}$
$S_5 = \{x \in \mathbb{N} \mid x > 71 \text{ and } x < 65\} = \phi$.

Except in S_5, which is empty, we can find in each of the subsets above a *least element*, i.e. an element of the subset which is less than or equal to every element of the subset in question. Thus the least elements of the subsets S_1, S_2, S_3 and S_4 are 0, 101, 6 and 0 respectively.

Having a least element is a common property, which is also shared by many other non-empty subsets of $\mathbb{N}$. For example, 0 is the least element of the entire $\mathbb{N}$, and the element a is the least element of the singleton subset $\{a\}$, which consists of a alone. In fact every set of natural numbers that we can think of will always have a least element as long as it is non-empty. Let us now formulate this simple property of natural numbers as:

2.2.1. The well-ordering principle. *Every non-empty set S of natural numbers has a unique least element, i.e., there is an element $a \in S$ such that $a \leqslant b$ for all elements $b \in S$.*

We shall accept this principle as a basic assumption for our subsequent discussion, because it cannot be derived from the usual laws and basic notion of natural numbers. In other words, we shall henceforth take as known and true and without question that every non-empty subset of $\mathbb{N}$ has a least element.

Finally it must be emphasized that the above principle is a property particular only to the natural numbers, and sets of other types of numbers (such as integers and rational numbers) may not have a least element.

Take for example the set $\mathbb{Z}$ of all integers which consists of all positive and negative whole numbers together with zero, 0, ± 1, ± 2, ... Again, given any two integers x and y, it is either $x < y$ or $x = y$ or $x > y$. Consider the following subsets of $\mathbb{Z}$:

$T_1 = \{-7, 9, -1, -5, 100\}$
$T_2 = \{13, 2, -5, -7, 0\}$
$T_3 = \{x \in \mathbb{Z} \mid x \geqslant 0\} = \mathbb{N}$
$T_4 = \{x \in \mathbb{Z} \mid x \leqslant 0\}$
$T_5 = \{0, 2, -2, 4, -4, \ldots\}$.

The sets T_1, T_2 and T_3 has least elements, which are -5, -7 and 0 respectively, while neither T_4 nor T_5 has a least element.

Similarly it is easy to see that not every set of rational numbers has a least element.

2.3. THE PRINCIPLE OF MATHEMATICAL INDUCTION

Imagine that an infinite row of dominoes was set up on edge close enough to one another, so that when anyone falls, it knocks down the next one in line. If the first one is pushed over, would all the dominoes fall down one after another? The statement below is an idealization of this notion:

Let S be a set of natural numbers with the following two properties:

(i) *0 belongs to S,*

(ii) *whenever a natural number k belongs to S, then the next number $k + 1$ also belongs to S.*

Then S is the set $\mathbb{N}$ of all natural numbers.

Condition (i) corresponds to the requirement that the first domino is being pushed over while condition (ii) corresponds to the requirement that the dominoes are close to one another.

It is on the basis of this principle that we were able to complete the second proof of the formula for the sum of an arithmetic progression given in Section 2.1. Let us now prove this principle using the well-ordering principle of Section 2.2.

Proof. Let S be a set of natural numbers that has the properties (i) and (ii) of the above statement and let T be the set of natural numbers which do not belong to S. Then either $T = \phi$ or T is a non-empty set. In the former case we have $S = \mathbb{N}$ and there is nothing more to be proved. Suppose now that T is non-empty, then by the well-ordering principle, T has a least element a, which, by definition of T, is not an element of S. By (i), $a \neq 0$, i.e. $a \geqslant 1$. Therefore $a - 1$ is a natural number. Then either $a - 1$ belongs to T or $a - 1$ does not belong to T. We shall see that either case will lead to a contradiction, thus showing the impossibility of the assumption that $T \neq \phi$. In the former case, $a - 1 \in T$ and $a - 1 < a$; this contradicts that $a \in T$ is the least element of T. In the latter case, $a - 1 \notin T$, we would have $a - 1 \in S$. Now it follows from $a - 1 \in S$ and (ii) that $(a - 1) + 1 = a$ is an

element of S, which contradicts $a \in T$. Therefore it is not possible that T is non-empty. Therefore $T = \phi$ and hence $S = \mathbb{N}$. The proof is now complete. ■

It is an easy exercise to dress up the second proof of Section 2.1 in the notation of the principle of mathematical induction.

2.3.1. *Example.* Prove that $0 + 1 + 2 + \ldots + n = \frac{n}{2}(n + 1)$ holds for all natural numbers n.

Proof. Let S be the set of natural numbers n for which the formula

$$0 + 1 + 2 + \ldots + n = \frac{n}{2}(n + 1)$$

holds: $S = \{n \in \mathbb{N} \mid 0 + 1 + 2 \ldots + n = \frac{n}{2}(n + 1)\}$. We want to prove that $S = \mathbb{N}$. For this purpose we need only show that conditions (i) and (ii) are satisfied. Condition (i) is trivial. For (ii), let us suppose that $k \in S$. Then

$$0 + 1 + 2 + \ldots + k + (k + 1) = \frac{k}{2}(k + 1) + (k + 1)$$

$$= \frac{(k + 1)}{2}(k + 2).$$

Therefore $(k + 1) \in S$. Hence $S = \mathbb{N}$ by the principle of mathematical induction. ■

Let us try to prove a more complicated formula by the same method.

2.3.2. *Example.* Show that for any natural number n

$$1^2 + 2^2 + 3^3 + \ldots + n^2 = \frac{1}{6}n(n + 1)(2n + 1).$$

The formula in question is actually the general expression of a collection of statements, one for each natural number:

$F(0)$: $0^2 = 0$

$F(1)$: $1^2 = 1$

$F(2)$: $1^2 + 2^2 = 5$

$F(3)$: $1^2 + 2^2 + 3^2 = 14$

. . . . . .

$F(k)$: $1^2 + 2^2 + \ldots + k^2 = \frac{1}{6}k(k + 1)(2k + 1)$

. . . . . .

We are therefore required to prove that all these statements are true. Naturally it is impossible for us to prove each statement one by one, and the most appropriate method of proof is an application of the principle of mathematical induction.

Proof. Let S be the set all natural numbers n, for which the corresponding statement $F(n)$ is true:

$$S = \{n \in \mathbb{N} \mid F(n) \text{ is true}\}$$

The aim of the proof is to show that $S = \mathbb{N}$. Thus it is sufficient to prove that the set S has properties (i) and (ii). Obviously S has property (i) because $F(0)$ is true by direct verification. To show that S has property (ii) we prove that $F(k+1)$ holds (i.e. $k+1$ belongs to S) under the assumption that $F(k)$ is true (i.e. $k \in S$). Let $F(k)$ be true. Then

$$\begin{aligned}
&1^2 + 2^2 + \ldots + k^2 + (k+1)^2 = [1^2 + 2^2 + \ldots + k^2] + (k+1)^2 \\
&= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2 \\
&= \frac{1}{6}(k+1)[k(2k+1) + 6(k+1)] \\
&= \frac{1}{6}(k+1)[(k+1)+1][2(k+1)+1].
\end{aligned}$$

Therefore $F(k+1)$ is true and hence S has property (ii). By the principle of mathemetical induction, $S = \mathbb{N}$,

i.e. $$F(n): 1^2 + 2^2 + \ldots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$ ∎

2.3.3. *Example.* Let $p_0, p_1, \ldots$ and $q_0, q_1, \ldots$ be two sequences of positive numbers such that

$$p_0 > q_0;\ p_{n+1} = \frac{1}{2}(p_n + q_n);\ \sqrt{p_n q_n} = c,\ n = 0, 1, 2, \ldots$$

where c is a positive constant. Prove that

$$p_n > p_{n+1} > c > q_{n+1} > q_n,\ n = 0, 1, 2, \ldots$$

Proof. We shall use the well-known inequality of arithmetic and geometric means: for any positive numbers x, y,

$$\sqrt{xy} \leqslant \frac{1}{2}(x+y)$$

and equality holds if and only if $x = y$. This follows from the simple observation that

$$0 \leqslant (\sqrt{x} - \sqrt{y})^2 = x - 2\sqrt{xy} + y$$

We are required to prove four separate inequalities:

(a) $p_n > p_{n+1}$; (b) $p_{n+1} > c$; (c) $c > q_{n+1}$; (d) $q_{n+1} > q_n$;

and we propose to do these by induction.

Basis of induction at $n = 0$.

(a) $p_0 - p_1 = p_0 - \frac{1}{2}(p_0 + q_0) = \frac{1}{2}(p_0 - q_0) > 0$, since $p_0 > q_0$.
Therefore $p_0 > p_1$.

(b) $p_1 - c = \frac{1}{2}(p_0 + q_0) - c > \sqrt{p_0 q_0} - c = 0$, by the inequality of the means. Therefore $p_1 > c$.

(c) $c - q_1 = c - c^2/p_1 = c(1 - c/p_1) > 0$, since $1 > c/p_1$ by (b).
Therefore $c > q_1$.

(d) $q_1 - q_0 = c^2/p_1 - c^2/p_0 = c^2(p_0 - p_1)/p_0 p_1 > 0$ by (a).
Therefore $q_1 > q_0$.

Induction step : Under the assumption that for some $k \geqslant 0$

$$p_k > p_{k+1} > c > q_{k+1} > q_k,$$

to prove that

$$p_{k+1} > p_{k+2} > c > q_{k+2} > q_{k+1}$$

(a) $p_{k+1} - p_{k+2} = p_{k+1} - \frac{1}{2}(p_{k+1} + q_{k+1}) = \frac{1}{2}(p_{k+1} - q_{k+1}) > 0$,
since $p_{k+1} > q_{k+1}$ by induction assumption. Therefore $p_{k+1} > p_{k+2}$.

(b) $p_{k+2} - c = \frac{1}{2}(p_{k+1} + q_{k+1}) - c > \sqrt{p_{k+1} q_{k+1}} - c = 0$, by the inequality of the means. Therefore $p_{k+2} > c$.

(c) $c - q_{k+2} = c - c^2/p_{k+2} = c(1 - c/p_{k+2}) > 0$ since $1 > c/p_{k+2}$ by (b).
Therefore $c > q_{k+2}$.

(d) $q_{k+2} - q_{k+1} = c^2/p_{k+2} - c^2/p_{k+1}$
$= c^2(p_{k+1} - p_{k+2})/p_{k+1} p_{k+2} > 0$ by (a).
Therefore $q_{k+2} > q_{k+1}$.

Putting (a), (b), (c) and (d) together we obtain, by the principle of mathematical induction, that

$$p_n > p_{n+1} > c > q_{n+1} > q_n$$

holds for all natural numbers n. ∎

It follows from the above examples that for certain application of the principle of mathematical induction, the following formulation of the principle may be particularly useful.

2.3.4. Principle of mathematical induction. *Let for each $n \in \mathbb{N}$, $P(n)$ be a statement about the natural number n. If the following two conditions are satisfied:*

(i) $P(0)$ is true;

(ii) if $P(k)$ is true, then $P(k+1)$ is true;

then $P(n)$ is true for every $n \in \mathbb{N}$.

2.3.5. *Example.* Prove that for any $x \neq 1$,

$$1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Proof. The statement is trivially true for $n = 0$. Suppose that it is true for $n = k$. Then

$$1 + x + \ldots + x^k + x^{k+1} = \frac{1 - x^{k+1}}{1 - x} + x^{k+1}$$

$$= \frac{1 - x^{k+2}}{1 - x}.$$

Therefore the statement for $n = k + 1$ is also true. Hence it is true by induction for any natural number n. ∎

2.4. EXERCISE

1. Examine each of the following sets of numbers and find out if it has a least element. If your answer is yes, exhibit this element. If your answer is no, show that given any element x of the set, you can always find another element y of the same set, such that $y < x$.

   (a) $\{x \in \mathbb{R} \mid x = \sin\theta \text{ for some } \theta \in \mathbb{R}\}$, where $\mathbb{R}$ denotes the set of all real numbers.

(b) $\{x \in \mathbb{N} \mid x \text{ is a prime number}\}$.

(c) $\{x \in \mathbb{R} \mid x = \frac{1}{n} \text{ for some } n \in \mathbb{N}\}$.

2. If p and q are any positive integers, prove that there exists a positive integer n such that $np \geqslant q$ (*Archimedean Property*).

3. Prove that $0^3 + 1^3 + 2^3 + 3^3 + \ldots + n^3 = \left[\frac{n(n+1)}{2}\right]^2$ for any natural number n.

4. Prove that $0 + 1 + 3 + 6 + \ldots + \frac{n(n+1)}{2} = \frac{n(n+1)(n+2)}{6}$ for any natural number n.

5. Prove that $\frac{0}{2^0} + \frac{1}{2^1} + \frac{2}{2^2} + \ldots + \frac{n}{2^n} = 2 - \frac{n+2}{2^n}$ for any natural number n.

6. Prove that $3^{4n+2} + 2^{6n+3}$ is divisible by 17 for any natural number n.

7. Prove that $(3n + 1)7^n - 1$ is divisible by 9 for any natural number n.

8. Prove that $3^{2n} - 32n^2 + 24n - 1$ is divisible by 512 for any natural number n.

9. (a) Prove that the product of any three consecutive natural numbers is divisible by 6.
   (b) Using (a) or otherwise, prove that the product of any four consecutive natural numbers is divisible by 24.

10. If a set has n elements, where n is any natural number, prove that the set has 2^n subsets.

2.5. MISCELLANEOUS REMARKS

We wish to emphasize that in a proof by induction, both conditions of the principle must be verified. The verification of condition (i) is called the *basis of induction* while that of condition (ii) is called the *induction step*.

The assumption (i.e. $k \in S$ or $P(k)$ holds) under which the induction step is carried out is called the *induction assumption*. A proof by induction cannot be deemed to be complete if either one of the two parts is missing. This can be compared with the notion of a row of dominoes. If either the first domino is not pushed over (i.e. there is no basis for induction), or if some gaps between the dominoes are too large (i.e. the induction step fails), then the complete row of dominoes will not fall down. To illustrate this point, we consider the two examples below:

2.5.1. *Example.* The statement

$$1 + 3 + 5 + \ldots + (2n - 1) = n^2 + 3 \text{ for all } n \in \mathbb{N}$$

is obviously false. Nevertheless we can show that it holds for $n = k + 1$ under the assumption that it holds for $n = k$. This is an example in which the induction step would be carried through while there is no basis for induction.

2.5.2. *Example.* The obviously absurd statement

$$n = 0 \text{ for all } n \in \mathbb{N}$$

is true for $n = 0$. Here we have a basis for induction but the induction step would fail.

Finally let us go through the following example, in which a hidden careless mistake leads to a blatantly absurd conclusion. We denote by $\max(a, b)$ the larger one of the two numbers a and b (i.e. $\max(a, b) = a$ if $a \geqslant b$ and $\max(a, b) = b$ if $a \leqslant b$).

2.5.3. *Example.* Go through the following very carefully to find the hidden mistake:

Fake theorem. $7 = 8$

Fake proof. Consider for each natural number n the statement

$F(n)$: If a and b are natural numbers such that $\max(a, b) = n$, then $a = b$.

We propose to 'prove' that $F(n)$ holds for every $n \in \mathbb{N}$ by induction.

(i) *Basis of induction.* Let a and b be natural numbers such that $\max(a, b) = 0$. Then $a \leqslant 0$ and $b \leqslant 0$. Therefore $a = 0$ and $b = 0$, because a and b are natural numbers. Hence $a = b$.

(ii) *Induction Step.* Suppose that $F(k)$ holds for some natural number k. Let a and b be two natural numbers such that $\max(a, b) = k + 1$. Backtracking one step from a and b we get $\max(a - 1, b - 1) = k$. By the induction assumption that $F(k)$ holds we obtain $a - 1 = b - 1$. Hence $a = b$. Therefore $F(k + 1)$ holds.

This completes the 'proof' by induction that $F(n)$ holds for every $n \in \mathbb{N}$. Take $n = 8$. Then it follows from $\max(7, 8) = 8$, and $F(8)$ being true that $7 = 8$. ∎

Where is the fault? It must be in the inductive 'proof' of the statement $F(n)$! It is not in part (i) because all arguments there are sound, so it must be in part (ii). In particular the fault must occur at the instance in which the induction assumption is applied. The induction assumption in this case should read: if $a - 1$ and $b - 1$ are natural numbers such that $\max(a - 1, b - 1) = k$, then $a - 1 = b - 1$. It is true that $\max(a - 1, b - 1) = k$, but by going back one step from the natural numbers a and b at the same time, $a - 1$ or $b - 1$ may not be a natural number anymore. For example, when a or b equals 0.

Therefore we are in no position to conclude that $a - 1 = b - 1$. Thus we see that the fault of the 'proof' is to take one step backward and out of the set $\mathbb{N}$!

2.6. ANOTHER VERSION OF THE PRINCIPLE OF MATHEMATICAL INDUCTION

Take the infinite row of dominoes again. Let r be a natural number. Suppose that the dominoes are spaced out in such a way that disregarding the first r dominoes (i.e. the 0th, 1st, 2nd, . . . , and $(r - 1)$th), anyone of them will fall down if all the preceeding ones (again disregarding the 0th, 1st, . . . , $(r - 1)$th) are knocked down. If the r-th domino is pushed over, we expect that beginning with the r-th one, all the dominoes will fall down. This state of affairs can be formulated as a variant of the original principle of mathematical induction. Since it can be shown to be actually equivalent to the original principle, we prefer to call it by the same name.

2.6.1. Principle of mathematical induction. *For each $n \in \mathbb{N}$, let $P(n)$ be a*

statement about the natural number n, and r be a natural number. Suppose that

(i) $P(r)$ *is true*

(ii) If for any $k \geqslant r$, *the statements* $P(r), P(r+1), \ldots, P(k-1), P(k)$ *are all true, then* $P(k+1)$ *is true.*

Then $P(n)$ *is true for all* $n \geqslant r$.

The proof of this version of the principle is completely similar to the one given to the original version in Section 2.3. This, as well as the proof of equivalence of the two versions of the principle, shall be left as an exercise to the interested students.

The present version of the principle differs from the original version in two aspects: (a) its starting point of induction is at $P(r)$ instead of at $P(0)$, (b) it has a stronger induction assumption. Because of (a), it is particularly useful for problems in which some of statements among $P(0), P(1), \ldots, P(r-1)$ are uninteresting, or meaningless. Secondly, under a stronger induction assumption, it is often easier to carry out the induction step.

2.6.2. *Example.* We have seen in Example 2.3.3 that given any two positive numbers x and y the geometric mean $\sqrt{xy}$ is always less than or equal to the arithmetic mean $(x+y)/2$, i.e. $\sqrt{xy} \leqslant (x+y)/2$. Here we shall try to obtain a generalized inequality.

Let $x_1, x_2, \ldots$ be a sequence of positive numbers, we define

$$A_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

and $$G_n = (x_1 x_2 \ldots x_n)^{1/n}$$

as the arithmetic mean and the geometric mean of the numbers $x_1, x_2, \ldots, x_n$ respectively. Prove that *the geometric mean of any n positive numbers is less than or equal to the arithmetic mean of the same n positive numbers,* i.e. $P(n)$: $G_n \leqslant A_n$ *holds for* $n = 2, 3, \ldots$ *and for any positive numbers* $x_1, x_2, \ldots$

Proof. The presence of the nth root in the expression of G_n makes matters very complicated and awkward for us to carry out the induction step in a straight-forward proof by induction. In order to avoid this difficulty, we shall proceed in two parts. Initially we shall prove by induction that $G_n \leqslant A_n$ holds for all the special values of n where $n = 2, 4, 8, \ldots, 2^p, \ldots$, being a power of 2. Then we shall take care of all the

other values of n lying in between any two consecutive powers of 2.

Part A. For all $n = 2, 4, \ldots, 2^p, 2^{p+1}, \ldots$

$$G_n \leqslant A_n$$

holds for any positive numbers $x_1, x_2, x_3, \ldots$.

Basis of induction at $p = 1$. Here

$$G_2 \leqslant A_2 : \sqrt{x_1 x_2} \leqslant \frac{1}{2}(x_1 + x_2)$$

is just the inequality of the two means which has been seen to hold in Example 2.3.3.

Induction step. Assume that $G_n \leqslant A_n$ holds for the k values of $n = 2, 2^2, \ldots, 2^k$ and for any $m = 2^k$ positive numbers $x_1, x_2, \ldots, x_m$. Then

$$(x_1 x_2 \ldots x_m)^{1/m} \leqslant \frac{1}{m}(x_1 + y_2 + \ldots + x_m).$$

Choosing another set of m positive numbers $x_{m+1}, x_{m+2}, \ldots, x_{2m}$, we also have

$$(x_{m+1}\, x_{m+2} \ldots x_{2m})^{1/m} \leqslant \frac{1}{m}(x_{m+1} + x_{m+2} + \ldots + x_{2m}).$$

It follows from these two inequalities and the inequality of the induction assumption at $n = 2$, we get

$$\begin{aligned}
&(x_1\, x_2 \ldots x_m\, x_{m+1} \ldots x_{2m})^{1/2m} \\
&= \{(x_1\, x_2 \ldots x_m)^{1/m}\, (x_{m+1}\, x_{m+2} \ldots x_{2m})^{1/m}\}^{1/2} \\
&\leqslant \frac{1}{2}\{(x_1\, x_2 \ldots x_m)^{1/m} + (x_{m+1}\, x_{m+2} \ldots x_{2m})^{1/m}\} \\
&\leqslant \frac{1}{2}\{\frac{1}{m}(x_1 + x_2 + \ldots + x_m) + \frac{1}{m}(x_{m+1} + x_{m+2} + \ldots x_{2m})\} \\
&= \frac{1}{2m}(x_1 + x_2 + \ldots + x_m + x_{m+1} + \ldots + x_{2m}).
\end{aligned}$$

Thus

$$(x_1\, x_2 \ldots x_{2m})^{1/2m} \leqslant \frac{1}{2m}(x_1 + x_2 + \ldots + x_{2m})$$

where $2m = 2^{k+1}$. The induction proof of Part A is now complete.

Part B. Given any natural number n, we can find two consecutive numbers $p - 1$ and p such that $2^{p-1} < n \leqslant 2^p$, i.e. n lies between two consecutive

powers of 2. This means that by back-tracking one step at a time, we can start at 2^p and reach n in less than 2^{p-1} steps. Correspondingly beginning with the statement $P(2^k)$, which has been shown to be true, we shall arrive at the statement $P(2^k - 1)$ by back-tracking one step. Therefore in a finite number of steps we shall reach the statement $P(n)$ which we want to prove. Obviously all that we need to do is to prove that one step backward from a valid statement $P(r)$ gives another valid statement $P(r-1)$. In other words, we have just to carry out an induction step backward, i.e. to prove that

$$(x_1\, x_2 \ldots x_{r-1})^{1/(r-1)} \leqslant \frac{1}{r-1}\,(x_1 + x_2 + \ldots + x_{r-1})$$

holds under the assumption that

$$(x_1\, x_2 \ldots x_r)^{1/r} \leqslant \frac{1}{r}(x_1 + x_2 + \ldots + x_r)$$

holds for any r positive numbers $x_1, x_2, \ldots, x_r$.

Put $G = (x_1\, x_2 \ldots x_{r-1})^{1/(r-1)}$ and $A = \dfrac{1}{r-1}\,(x_1 + x_2 + \ldots + x_{r-1})$.

Then by our assumption we get

$$(x_1 x_2 \ldots x_{r-1} G)^{1/r} \leqslant \frac{1}{r}(x_1 + x_2 + \ldots + x_{r-1} + G),$$

$$(G^{r-1} G)^{1/r} \leqslant \frac{1}{r}[(r-1)A + G],$$

$$rG \leqslant (r-1)A + G.$$

Since $r \geqslant 2$, we obtain $G \leqslant A$, i.e.

$$(x_1\, x_2 \ldots x_{r-1})^{1/(r-1)} \leqslant \frac{1}{r-1}\,(x_1 + x_2 + \ldots + x_{r-1}) \qquad \blacksquare$$

2.7. EXERCISE

1. Prove that $1^2 + 3^2 + 5^2 + \ldots + (2n-1)^2 = \dfrac{n(2n-1)(2n+1)}{3}$ for any natural number $n \geqslant 1$.

2. Prove that $1^3 + 3^3 + 5^3 + \ldots + (2n-1)^3 = n^2(2n^2 - 1)$ for any natural number $n \geqslant 1$.

3. Prove that $1 \times n + 2(n-1) + 3(n-2) + \ldots + (n-1)2 + n \times 1 = \dfrac{n(n+1)(n+2)}{6}$ for any natural number $n \geqslant 1$.

4. Prove that $\frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \ldots + \frac{1}{n(n+1)} = \frac{n}{n+1}$ for any natural number $n \geqslant 1$.

5. Prove that $\frac{1}{1 \times 2 \times 3} + \frac{1}{2 \times 3 \times 4} + \ldots + \frac{1}{n(n+1)(n+2)} = \frac{n(n+3)}{4(n+1)(n+2)}$ for any natural number $n \geqslant 1$.

6. Prove that
$(n+1)(n+2)(n+3) \ldots (2n) = 2^n \times 1 \times 3 \times 5 \times \ldots \times (2n-1)$
for any natural number $n \geqslant 1$.

7. Prove that $(1 - \frac{1}{2^2})(1 - \frac{1}{3^2})(1 - \frac{1}{4^2}) \ldots (1 - \frac{1}{n^2}) = \frac{n+1}{2n}$ for any natural number $n \geqslant 2$.

8. Prove that $(1 - \frac{4}{1^2})(1 - \frac{4}{3^2})(1 - \frac{4}{5^2}) \ldots [1 - \frac{4}{(2n-1)^2}] = \frac{2n+1}{1-2n}$ for any natural number $n \geqslant 1$.

9. Prove that $6^n + 4$ is divisible by 5 for any natural number $n \geqslant 1$.

10. Prove that, for any natural number $n \geqslant 1$, $\frac{2^n - (-1)^n}{3}$ is an odd number.

11. Prove that $2^n > n^2$ for any natural number $n \geqslant 5$.

12. If $x > -1$, prove that $(1+x)^n \geqslant 1 + nx$ for any natural number $n \geqslant 1$. (Bernoulli's inequality)

*13. If $a > 0$, prove that
$$a^n + a^{n-2} + a^{n-4} + \ldots + \frac{1}{a^{n-4}} + \frac{1}{a^{n-2}} + \frac{1}{a^n} \geqslant n + 1$$
for any natural number $n \geqslant 1$.

*14. Prove that $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \ldots + \frac{1}{\sqrt{n}} \geqslant \sqrt{n}$ for any natural number $n \geqslant 1$.

*15. Prove that $\frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots + \frac{1}{n^2} < \frac{n-1}{n}$ for any natural number $n \geqslant 2$.

*16. Prove that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3} + \ldots + \frac{1}{2n} \geqslant \frac{7}{12}$$

for any natural number $n \geqslant 2$.

17. Prove that $a^n - 1 = (a-1)(a^{n-1} + a^{n-2} + a^{n-3} + \ldots + a + 1)$ for any natural number $n \geqslant 1$.
(Hint: $a^{n+1} - 1 = (a+1)(a^n - 1) - a(a^{n-1} - 1)$)

18. Suppose there are n ($\geqslant 2$) numbers, each of which is equal to the sum of the squares of two integers. Prove that the product of these n numbers is equal to the sum of the squares of two integers.

19. There are n ($\geqslant 1$) straight lines on a plane. No two lines are parallel and no three are concurrent. Prove that these n straight lines
    (a) have $\frac{n(n-1)}{2}$ points of intersection,
    (b) divide themselves into n^2 line segments,
    (c) divide the plane into $1 + \frac{n(n+1)}{2}$ regions.

20. For any natural number $n \geqslant 8$, prove that
$$n = 3p + 5q$$
for some natural numbers p and q.

*21. Let $u = x + y$ and $v = xy$. For any natural number n, prove that $x^n + y^n$ can be expressed as a polynomial in u and v.

*22. Prove that $n^{n+1} > (n+1)^n$ for any natural number $n \geqslant 3$.

*23. Prove that $2^n > n^3$ for any natural number $n \geqslant 10$.

*24. Prove that $\sqrt[n]{n} < \sqrt[3]{3}$ for any natural number $n \geqslant 4$.

*25. Prove that $(\sqrt{3}+1)^{2n+1} - (\sqrt{3}-1)^{2n+1}$ is equal to an integer which is divisible by 2^{n+1} for any natural number n.

*26. If n is a natural number, prove that

$$\frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n \cdot \sqrt{5}}$$

is also a natural number.

*27. What is wrong with the following proof that all cars have the same colour?

Proof. Let n be the number of cars. When $n = 1$, the statement is clearly true, that is, one car has the same colour, whatever colour it is. Assume that any group of n cars have the same colour. Now consider a group of $(n + 1)$ cars. Taking any n of them, the induction hypothesis states that they all have the same colour, say red. The only issue is the colour of the remaining 'uncoloured' car. Consider, therefore, any other group of n of the $(n + 1)$ cars that contain the uncoloured car. Again, by the induction hypothesis, all the cars in the new group must have the same colour. Then, since all of the coloured cars in this group are red, the uncoloured car must also be red.

2.8. RECURSIVE FORMULAE

We have seen in the previous sections that the principle of mathematical induction has provided us with the justification of establishing an infinite sequence of statements step by step in a precisely prescribed manner. The same principle can serve as a base for laying down an infinite sequence of definitions in a similar manner. This can be illustrated by the definition of factorials $n!$ which is in fact an infinite sequence of symbols, $0!$, $1!$, $2!$, $\ldots$, $n!$, $\ldots$ one for each natural number n, defined by the conditions

(a) $0! = 1$

(b) $(k + 1)! = (k + 1)(k!)$.

The statement (a) puts down the *initial* meaning of the symbol $n!$ at $n = 0$ by defining $0!$ to be the number 1. In (b) we have a *recursive formula* that instructs us how to work out $(k + 1)!$ on the assumption that $k!$ is already defined. Thus proceeding step by step we have

$$0! = 1$$
$$1! = 1 \cdot 0! = 1$$
$$2! = 2 \cdot 1! = 2 \cdot 1 = 2$$
$$3! = 3 \cdot 2! = 3 \cdot 2 \cdot 1 = 6$$
$$\cdots\cdots\cdots$$

By virtue of the principle of mathematical induction, each symbol of the entire sequence $0!, 1!, 2!, \ldots, n!, \ldots$ has now an unambiguous meaning.

2.8.1. *Example.* In a similar way the convenient summation symbol Σ and product symbol Π are defined. Let $a_1, a_2, \ldots, a_n, \ldots$ be an infinite sequence of numbers. Then for any natural number $n \geqslant 1$, the sum $\Sigma_{i=1}^{n} a_i$ and the product $\Pi_{i=1}^{n} a_i$ are defined on the basis of the principle of mathematical induction by the following formulae:

(a) Initial formulae

$$\sum_{i=1}^{1} a_i = a_1; \quad \prod_{i=1}^{1} a_i = a_1$$

(b) Recursive formulae

$$\sum_{i=1}^{k+1} a_i = \left(\sum_{i=1}^{k} a_i\right) + a_{k+1}; \qquad \prod_{i=1}^{k+1} a_i = \left(\prod_{i=1}^{k} a_i\right) a_{k+1}$$

Thus

$$\sum_{i=1}^{1} a_i = a_1; \qquad \prod_{i=1}^{1} a_i = a_1$$
$$\sum_{i=1}^{2} a_i = a_1 + a_2; \qquad \prod_{i=1}^{2} a_i = a_1 a_2$$
$$\sum_{i=1}^{3} a_i = a_1 + a_2 + a_3; \qquad \prod_{i=1}^{3} a_i = a_1 a_2 a_3$$
$$\cdots\cdots\cdots\cdots \qquad \cdots\cdots\cdots\cdots$$

The letter i in the expression $\Sigma_{i=1}^{n} a_i$ is called the *summation index* while 1 and n are the *lower bound* and the *upper bound* of the summation. The letter for the summation index may be changed without affecting the sum as long as the bounds remain the same. Thus

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j.$$

Similar notations apply to the expression $\prod_{i=1}^{n} a_i$.

2.8.2 *Example.* The *Fibonacci numbers* $F(n)$ are defined inductively by

(a) $F(1) = F(2) = 1$

(b) $F(k+1) = F(k-1) + F(k)$.

Thus the sequence of Fibonacci numbers is

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

each term being the sum of the preceding two terms. This sequence of numbers appeared first in a book entitled *Liber Abaci*, written in the 13th century by Leonardo of Pisa, who also called himself Fibonacci, being the son of Bonacci.

The following problem is found in the book:

> 'A man puts one pair of rabbits in a certain place entirely surrounded by a wall. How many pairs of rabbits can be produced from that pair in a year if the nature of these rabbits is such that every month each pair bears a new pair, which in two months becomes productive?'

According to the text the maturing process is that a newborn pair become an adolescent pair in the following month and become productive adults another month later. Denote by $N(n)$, $Y(n)$, and $A(n)$ the number of newborn, adolescent and adult pairs in the n-th month, and assuming that none of the rabbits of the colony dies, then

$N(n) = A(n)$ (Every adult pair give one newborn pair every month.)

$Y(n) = N(n-1)$ (Newborn pairs become adolescents in 1 month.)

$A(n) = A(n-1) + Y(n-1)$ (Adult rabbits do not die and adolescents become adults in 1 month.)

Suppose that the original ancestral pair is a newborn pair, then the population profile of the colony is shown in the table below:

nth Month	$A(n)$	$Y(n)$	$N(n)$	Total $T(n)$
1st			1	1
2nd		1		1
3rd	1		1	2
4th	1	1	1	3
5th	2	1	2	5

nth month	$A(n)$	$y(n)$	$N(n)$	Total $T(n)$
6th	3	2	3	8
7th	5	3	5	13
8th	8	5	8	21
9th	13	8	13	34
10th	21	13	21	55
11th	34	21	34	89
12th	55	34	55	144

Alternatively we may calculate $T(n)$ from the three equations of the maturing process as follows:

(a) $T(1) = A(1) + Y(1) + N(1) = 0 + 0 + 1 = 1$
$T(2) = A(2) + Y(2) + N(2) = 0 + 1 + 0 = 1$

(b) $T(n) = A(n) + Y(n) + N(n)$
$= 2A(n) + Y(n) = 2[A(n-1) + Y(n-1)] + N(n-1)$
$= T(n-1) + A(n-1) + Y(n-1)$
$= T(n-1) + [A(n-2) + Y(n-2)] + N(n-2)$
$= T(n-1) + T(n-2).$

Thus the population of the rabbit conlony is given by

(a) $T(1) = T(2) = 1$

(b) $T(n) = T(n-1) + T(n-2)$

which are precisely the initial and the recursive formulae of the Fibonacci numbers. Therefore $T(n) = F(n)$.

2.8.3. *Example.* There are three kinds of bees, the drones (male), the queen bees (female) and the worker bees (female). They all hatch from eggs laid by the queen bees. While a male bee hatches from an unfertilized egg, a female bee hatches from a fertilized egg. Thus a drone has only one parent while a queen bee or a worker bee has two. Find the number of ancestors of a drone in the n-th generation, assuming no incest.

Solution. Denote by $M(n)$, $F(n)$ and $A(n)$ the number of n-th generation male ancestors, female ancestors and all ancestors.

Then $M(1) = 0, F(1) = 1; M(2) = 1, F(2) = 1.$

Also $M(n+1) = F(n) = M(n-1) + F(n-1) = A(n-1)$

and $F(n+1) = M(n) + F(n) = A(n).$

Therefore the total number of the n-th generation ancestors of a drone is given by

(a) $A(1) = 1, A(2) = 2$

(b) $A(n+1) = A(n) + A(n-1)$.

Comparing $A(n)$ with the Fibonacci $F(n)$, we see that $A(n) = F(n+1)$ for all $n = 1, 2, 3, \ldots$

2.9. EXERCISE

1. Prove that $n! > n^2$ for any natural number $n \geqslant 4$.

2. Prove that $n! > n^3$ for any natural number $n \geqslant 6$.

3. Prove that $\sum_{i=1}^{n} i(i!) = (n+1)! - 1$ for any natural number $n \geqslant 1$.

4. A sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is defined by

$$a_1 = \sqrt{6}$$

and $$a_{n+1} = \sqrt{6 + a_n} \quad (n = 1, 2, 3, \ldots).$$

Prove that $a_n < a_{n+1}$ for any natural number $n \geqslant 1$.

5. A sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is defined by

$$a_{n+1} = a \cdot a_n + b \quad (n = 1, 2, 3, \ldots)$$

where a and b are constants such that a is neither 0 nor 1.

Prove that $$a_n = a^{n-1} a_1 + \frac{b}{a-1}(a^{n-1} - 1)$$

for any natural number $n \geqslant 1$.

6. A sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is defined by

$$a_1 = 1$$

and $$a_{n+1} = a \cdot a_n + n + 1 \quad (n = 1, 2, 3, \ldots)$$

where a is a constant.

(a) Find an expression for a_n when $a = 1$.

(b) Show that, when $a \neq 1$, a_n is of the form

$$A a^n + B n + C$$

where A, B and C are constants. Express A, B and C in terms of a.

7. Being given that $a_1 = 1$

$$\text{and} \qquad a_{n-1}\, a_n = \frac{1}{2^{n(n-1)}} \frac{\prod_{i=n}^{2n-1} (i!)}{\prod_{i=1}^{n-1} (i!)} \qquad (n = 2, 3, 4, \ldots),$$

prove that $a_n = \prod_{i=1}^{n} (2i-1)^{n+1-i}$ for any natural number $n \geqslant 1$.

*8. Prove that $\dfrac{(2n)!}{(n!)^2} > \dfrac{4^n}{n+1}$ for any natural number $n \geqslant 2$.

*9. Consider the so-called *Lucas sequence*

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

defined by $a_1 = 1,$

$a_2 = 3$

and $a_n = a_{n-2} + a_{n-1} \quad (n = 3, 4, 5, \ldots)$

Prove that

(a) $a_n = \left(\dfrac{1+\sqrt{5}}{2}\right)^n + \left(\dfrac{1-\sqrt{5}}{2}\right)^n$

(b) $a_n < \left(\dfrac{7}{4}\right)^n$

for any natural number $n \geqslant 1$.

*10. A sequence $a_0, a_1, a_2, a_3, \ldots, a_n, \ldots$ satisfies the relation

$$2(n+2)\, a_{n+2} - 3n\, a_{n+1} + (n-1)\, a_n = 0$$

with $a_0 = 1$ and $a_1 = \dfrac{1}{2}$.

Prove that $a_n = \dfrac{1}{2^n}$ for any natural number $n \geqslant 1$.

*11. Let α and β be the distinct real roots of the equation

$$x^2 - px + q = 0, \text{ where } p \neq 1 \text{ and } p^2 > 4q.$$

A sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is defined by

$$a_1 = p - \frac{q}{p-1}$$

and

$$a_n = p - \frac{q}{a_{n-1}} \qquad (n = 2, 3, 4, \ldots).$$

Prove that $a_n = \dfrac{(\alpha^{n+2} - \beta^{n+2}) - (\alpha^{n+1} - \beta^{n+1})}{(\alpha^{n+1} - \beta^{n+1}) - (\alpha^n - \beta^n)}$

(Problems 12-20 refer to Fibonacci numbers.)

12. Prove that $F(m + n) = F(m - 1)\,F(n) + F(m)\,F(n + 1)$ for any natural numbers $m \geqslant 2$ and $n \geqslant 1$.
(Hint: For an arbitrary m carry out an induction on n.)

13. Prove that $\sum_{i=1}^{n} F(i)^2 = F(n)\,F(n + 1)$ for any natural number $n \geqslant 1$.
(Hint: For $n \geqslant 2$, $F(n)^2 = F(n)\,F(n + 1) - F(n)\,F(n - 1)$.)

14. Prove that $2\sum_{i=1}^{n} F(i)^2 + F(n - 1)^2 - F(n)^2 = F(n + 1)^2$ for any natural number $n \geqslant 3$.

15. Prove that $\sum_{i=1}^{n} i\,F(i) = (n + 1)\,F(n + 2) - F(n + 4) + 2$ for any natural number $n \geqslant 1$.

16. Prove that $\sum_{i=1}^{n} F(2i - 1) = F(2n)$ for any natural number $n \geqslant 1$.

17. Prove that $\sum_{i=1}^{n} F(2i) = F(2n + 1) - 1$ for any natural number $n \geqslant 1$.

18. Prove that $\sum_{i=1}^{n} (-1)^{i+1} F(i) = 1 + (-1)^{n+1} F(n - 1)$ for any natural number $n \geqslant 2$.

19. Prove that every positive integer can be represented as a finite sum of Fibonacci numbers, none used more than once.
(Hint: Show for $n > 2$, every positive integer less than $F(n)$ is a sum of some of the numbers $F(1), F(2), \ldots, F(n - 2)$, none repeated.)

20. Show that the Binet formula

$$F(n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1 + \sqrt{5}}{2}\right)^n - \left(\frac{1 - \sqrt{5}}{2}\right)^n\right]$$

holds for any natural number $n \geqslant 1$.

3. Combinatorics

In this chapter we are interested in counting the different ways in which a given event can occur. There is a great variety of such counting problems that we encounter in every day life and in mathematics. Here we shall discuss a few useful methods that can be applied to such problems.

3.1. BOXES AND BALLS

We begin with a simple but typical problem.

3.1.1. *Problem.* There are two balls, one red and one blue, and three boxes numbered 1, 2 and 3. Find the number of ways in which the two balls can be put in the boxes, if each box can hold no more than one ball.

Solution (a). With only two balls and three boxes the problem is easy enough to be solved experimentally. The six different ways are (Fig. 3.1):

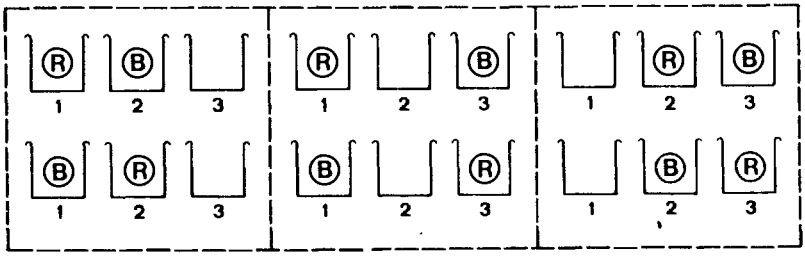


Fig. 3.1

They are arranged in three groups of two placements each with the same box empty. ■

In view of the fact that we shall be considering similar problems with far more balls and boxes, as well as variations of this problem in which some of the balls are not distinguishable from one another or the boxes

have different capacities, we shall solve this problem by an alternative method which can be applied to more complicated situations.

Solution (b). By an *event* we mean something that takes place or happens. Thus putting the two given balls into two of the three given boxes is an event, and we are interested in the number of distinct ways in which this event can occur. Let us denote this event by A and this number by $w(A)$. We can also think of event A as a sequence of two events B and C: event B is to put the blue ball into one of the given boxes and event C is to put the red ball into one of the two boxes which remain empty after event B has taken place. Thus event A occurs whenever event B and event C occur together and vice versa. Let us count the occurrence of the events. Event B can occur in 3 distinct ways, and event C can occur in 2 distinct ways after event B has occurred. Therefore

$$w(A) = w(B)\, w(C) = 3 \times 2 = 6. \qquad \blacksquare$$

In the very last step of the above solution (b) we have used a rather self-evident rule on the simultaneous occurrence of two events. Since we shall be making use of this rule frequently in our discussion, we formulate it as:

3.1.2. Rule of product. *If one event can occur in m ways and another event can occur in n ways, then there are $m \times n$ ways in which these two events can occur together.*

Using the method of solution (b), we can handle with ease any similar problem, with any number $n \geqslant 2$ of distinct boxes: there are altogether $n(n-1)$ ways to place two distinctly coloured balls in n distinctly numbered boxes.

Suppose we now have r distinguishable balls and n distinguishable boxes, where $r \leqslant n$. Then there are n ways to place the first ball in the n boxes, $n-1$ ways to place the second ball in the $n-1$ remaining empty boxes, $n-2$ ways to place the third ball in the now $n-2$ remaining empty boxes, and so forth. Applying the rule of product $r-1$ times, we obtain:

3.1.3. Theorem. *The total number of distinct ways to put r distinguishable balls in $n (n \geqslant r)$ distinguishable boxes is*

$$n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!}.$$

It is clear that with appropriate interpretation of the balls and boxes, Theorem 3.1.3 is adoptable to a great variety of problems of everyday life.

3.1.4. *Example.* A football team of 17 members is given 19 single rooms in a hostel. In how many ways can the members of the team be assigned to their rooms?

Solution. We can take the members of the team as balls and the rooms as boxes. Therefore the 17 members of the team can be accomodated in

$$\frac{19!}{2!} = 19 \times 18 \times \ldots \times 4 \times 3$$

ways in the 19 single rooms. ∎

3.1.5. *Example.* Three office bearers are to be selected among the students of Form Six, one from Form Upper Six to be the Chairman and two from Form Lower Six to be the Vice-Chairman and the Secretary. There are 21 students in the upper form and 25 students in the lower form. In how many ways can the office bearers be selected?

Solution. For the selection in Form Upper Six we use the model of 21 boxes (students) and 1 ball (position). Therefore we have 21 ways to select an office bearer. Similarly we have 25 x 24 ways to select two office bearers in Form Lower Six. By the rule of product we have 21 x 25 x 24 = 12,600 ways of selections. ∎

3.2. EXERCISE

1. There are 10 ships plying between two towns; in how many ways can a man go from one town to the other and return by a different ship?

2. Three travellers arrive at a town where there are four hotels; in how many ways can they take up their quarters, each at a different hotel?

3. How many different cubes with the six faces numbered from 1 to 6 can be made, if the sum of the numbers on each pair of opposite faces is 7?

4. How many different numbers of 5 digits may be formed with 0, 1, 2, 3, 4, 5, 6, 7? How many of these will be even?

*5. In a table-tennis tournament there are $2n$ participants. In the first round of the tournament each participant plays just once, so there are n games each occupying a pair of players. Find the number of different ways of arranging the pairing for the first round.

3.3. REMARKS

We have seen that in examining the occurrence of a complicated event, we first identify a number of events which are components that make up the given event, and then we apply the rule of product to get the required number. Obviously a selection of wrong events will lead to false results and care must be taken to make sure that when fitted together the selected components produce the given event. To illustrate our point, we take a second look at solution (b) to Problem 3.1.1. The event under consideration is

Event A: to put a blue ball and a red ball in boxes 1, 2, 3, no box holding more than one ball.

As components of Event A we have identified

Event B: to put the blue ball into one of the 3 boxes.

Event C: to put the red ball into one of the 2 remaining empty boxes (after the occurrence of Event B).

Events B and C together make up Event A because they take place together whenever Event A occurs and vice versa. We may also consider other components.

Event D: to put the red ball into one of the 3 boxes

Event E: to put the blue ball into one of the 2 remaining empty boxes (after the occurrence of Event D).

Similarly Events D and E together produce Event A and hence $w(A) = w(D)w(E) = 3 \times 2 = 6$. On the other hand the components B and D together will not produce Event A because of the requirement that no box should hold 2 balls.

3.4. PERMUTATIONS

3.4.1. *Problem.* There are n distinguishable balls and r distinguishable boxes where $n \geqslant r$. In how many ways can the boxes be filled by one ball each?

Solution. Clearly it makes no difference whether we fill all the boxes at the same time or we fill them one after another. Therefore, we view the event of filling the boxes simultaneously as composed of a series of component events. There are n ways to fill the first box, $n-1$ ways to fill the second box (because one ball is already put in the first box and only $n-1$ balls are left), $n-2$ ways to fill the third box and so forth. Applying the rule of product we obtain

$$n(n-1)(n-2)\ldots(n-r+1)=\frac{n!}{(n-r)!}$$

as the number of distinct ways to fill the boxes as prescribed by the problem. ∎

The above problem can be interpreted as the classical problem of arranging n distinct objects (e.g. letters A, B, . . .) taken r of them at a time, by thinking the balls as objects and the boxes as positions in the line. It is customary to call each such arrangement a *permutation of n objects taken r at a time.* For example, with $n=3$ and $r=2$, we have the six permutations

AB, AC, BA, BC, CA, CB

of the 3 objects A, B, C, taken 2 at a time. It is also customary to write

$$P(n,r)=n(n-1)(n-2)\ldots(n-r+1)=\frac{n!}{(n-r)!}$$

where $n \geqslant r$. Thus we have proved

3.4.2 Theorem. *There are*

$$P(n,r)=n(n-1)(n-2)\ldots(n-r+1)$$

permutations of n distinct objects taken r at a time, r being no greater than n.

3.4.3. *Example.* (a) Find the number of integers between 100 and 500 which are formed by the six digits 0, 2, 3, 4, 5, 6, no digit being used more than once. (b) How many of these integers are even?

Solution. (a) All the integers in question have 3 digital places. Therefore we may use the model of 3 boxes (digital places) and 6 balls (digits) with some slight modification. Because of the restriction on the range of the integers, only the digits 2, 3, 4 can be used for the hundredth digital place. Therefore we have 3 possibilities to fill the first box, then 5 for the second

and then 4 for the last. By the rule of product, the number of integers in question is $3 \times 5 \times 4 = 60$.

(b) To find the number of even integers, further modification to the model of boxes and balls has to be made. Now only the digits 0, 2, 4, 6 may be used for the unit digital place. In conjunction with the restriction to 2, 3, 4 for the hundredth digital place, we should distinguish two types of even integers within the range, type (i) being even integers with the digit 3 at the hundredth digital place, and type (ii) those with digit 2 or 4 at the hundredth digitial place. Let us count those of type (i). For the first box there is only 1 possibility (the digit 3). For the third box we have 4 possibilies (the digits 0, 2, 4 or 6) and the second box also 4 (the remainder of all six given digits). Therefore we have $1 \times 4 \times 4$ even integers of type (i). For those of type (ii) there are 2 choices (the digits 2 or 4) for the first box, 3 choices (the unused ones among 0, 2, 4, 6) for the third box and 4 choices (the remainder among all six given digits) for the second box. Therefore we have $2 \times 4 \times 3 = 24$ even integers of type (ii). Since each even integers in question is either of type (i) or of type (ii), the total number of such integers is

$$16 + 24 = 40. \qquad \blacksquare$$

In the very last step, we have made use of one instance of the following general rule.

3.4.4. Rule of sum. *If one event can occur m ways and another event can occur n ways, then there are $m + n$ ways in which exactly one of these two events can occur.*

3.4.5. *Example.* There are 22 different books, 5 on mathematics 7 on literature and 10 on natural science. The winner of a competition can select any two books of different subjects. After the winner has made his selection, the runner-up can choose any two books of different subjects from the remainder. In how many ways can the winner make his selection? Does the runner-up has just one choice less than the winner if the winner has selected one mathematics and one science book?

Solution. There are three subject combinations, Math.-Lit., Math.-Sci. and Lit.-Sci. The choices of the winner in each combination are as follows:

Math.-Lit.: $5 \times 7 = 35$

Math.-Sci.: $5 \times 10 = 50$

Lit.-Sci.: $7 \times 10 = 70$

By the rule of sum, the winner has $35 + 50 + 70 = 155$ ways to make his selection.

For the runner-up, there are 4 mathematics books, 7 literature books and 9 science books left. Therefore the number of choices is $4 \times 7 + 4 \times 9 + 7 \times 9 = 127$. In fact the runner-up has 28 choices less than the winner! ■

3.5. EXERCISE

1. Four persons enter a car in which there are six seats. In how many ways can they take their places?

2. There are nine different books on a shelf; four are red and five are green. In how many ways is it possible to arrange all the books on the shelf, if the colours must alternate?

3. There are two works each of three volumes, and two works each of two volumes. In how many ways can the ten books be placed on a shelf, so that volumes of the same work are not separated?

4. Calculate the number of arrangements, in which three men and two women can seat themselves in a railway carriage, designed to seat four on each side, if two of the four corner seats are to be occupied by women.

5. Find the number of ways of arranging n people in a straight row, if two particular people must always be separated.

6. Given an arrangement in a line of n people, find in how many ways they may be rearranged, if neither of the two assigned people is to occupy their original places.

7. It is required to arrange n people in a row so that three particular people do not all come together. Find the number of different arrangements.

3.6. PERMUTATIONS IN WHICH REPETITIONS ARE ALLOWED

Natural numbers less than 1000 can be regarded as permutations of the

10 digits 0, 1, 2, . . . , 9 taken 3 at a time, where each digit may appear repeatedly. For example, the number 010 is a permutation of the digits 0 and 1 where 0 appears twice, and in 222 the digit 2 appears thrice. There are 10^3 such numbers; therefore we have also 10^3 permutations. This suggests a variation of the boxes and balls problem as follows.

3.6.1. *Problem.* There are n distinguishable kinds of balls with an unlimited supply of each kind. There are r distinguishable boxes. In how many ways can we put one ball in each box.

Solution. We take one box at a time. In the first box we can put a ball of any given kind. Therefore there are n possibilities for the first box. Since there is further supply of the kind that has been used, there are still n possibilities for the second box and so also for all other boxes. By the rule of product we can do this in $n \times n \times \ldots \times n = n^r$ distinct ways. Therefore we have n^r ways to put one ball in each of the r distinguishable boxes if there are n different kinds of balls with an unlimited supply of each kind. ∎

Putting it in the customary language, we have the following theorem.

3.6.2. Theorem. *There are n^r permutations of n distinct objects taken r at a time, in which the objects may appear any number of times.*

3.6.3. *Example.* Consider all permutations of the letters A, B, C, D, E, taken 5 at a time with repetitions allowed. If they were written down in lexicographic order, at what place would you find the permutation DBCAC?

Solution. A permutation which precedes DBCAC must be exactly one of the following types:

(i) permutations where the first letter is either A, B or C (A _ _ _ _, B _ _ _ _ or C _ _ _ _)

(ii) permutations where the first letter is D and the second letter is A (DA _ _ _)

(iii) permutations where the first two letters are written DB and the third letter is either A or B (DBA _ _ or DBB _ _)

(iv) permutations where the first four letters are written DBCA and the fifth letter is either A or B (DBCAA or DBCAB)

Therefore the numbers of these kinds of permutations are respectively

$$3 \times 5^4, \quad 5^3, \quad 2 \times 5^2, \quad \text{and } 2.$$

By the rule of sum, there are $3 \times 5^4 + 5^3 + 2 \times 5^2 + 2 = 2{,}052$ permutations preceding DBCAC in the list. Hence DBCAC is at the 2,053rd place in the list. ■

3.6.4. *Example.* Given a set X of n elements. How many subsets of X are there?

Solution. Each subset of X is a selection of elements of X; different selections give rise to different subsets. We propose to count all such selections. The elements of X are thought as boxes, so we have n distinct boxes. Take two kinds of balls, one kind labelled with the letter S (selected) the other kind with R (rejected). Take an unlimited supply (actually n pieces would be sufficient) of each kind. Then we put one ball in each box. Take all those elements of X corresponding to boxes with an S ball to form a subset of X. In this way the number of subsets of X is the same as the number of different placements. With n boxes and 2 kinds of balls the number is 2^n. Therefore there are 2^n subsets of X. ■

3.7. EXERCISE

1. On three different days, a woman has to drive to a railway station in one of her five different cars. In how many ways can she make the three journeys?

2. In how many ways can five different prizes be given to four girls, when each girl is eligible for all the prizes?

3. (a) How many integers between 100 and 999 inclusive have distinct digits?
   (b) Of the integers with distinct digits in (a), how many are odd numbers?

4. How many positive integers are there having four digits, each digit being 1, 2, 3 or 4? How many of them have two or more equal digits?

5. Find the number of permutations of n different objects taken from 1 to r times when any of the objects may appear repeatedly.

6. Show that the number of permutations of n different things taken r at a time, when repetitions are allowed, but no consecutive repetitions is $n(n-1)^{r-1}$.

7. Find the sum of all the numbers of not more than 3 digits that can be formed with the digits 3, 2 and 1.

*8. Let $A = \{x \in \mathbb{N} : 1 \leqslant x \leqslant n\}$ where n is a positive integer and $\mathbb{N}$ is the set of all natural numbers.
(a) How many subsets of A contain at least one even integer?
(b) How many subsets of A contain exactly one even integer?
(Hint: consider separately the cases when n is even, and when n is odd.)

3.8. PERMUTATIONS OF OBJECTS SOME OF WHICH ARE ALIKE

Take four pieces from a game of scrable; two pieces bear the same letter A, one B and one C. There are twelve possible arrangements of these pieces. When listed in lexicographic order, they are:

AABC AACB ABAC ABCA ACAB ACBA
BAAC BACA BCAA CAAB CABA CBAA

Compared with the 24 possible arrangements of four different objects, the reduction in the number of arrangements by a factor of 2 must be due to the fact that 2 of the given pieces are alike and indistinguishable. This reduction factor can be explained and accounted for by the following experiment. Label the two like pieces with the numbers 1 and 2. Arrange these four now distinguishable pieces in 4! = 24 ways. Group these 24 permutations in pairs in which the pieces B and C have identical positions, for example,

$$A_1BA_2C \text{ and } A_2BA_1C$$
$$CA_1A_2B \text{ and } CA_2A_1B$$

After taking off the numbered labels, the two arrangements in each pair become the same permutation on the original lexicographic list. For example, A_1BA_2C and A_2BA_1C will be the permutation ABAC on the list while CA_1A_2B and CA_2A_1B yield the permutation CAAB. Therefore the 24 permutations of the 4 distinguishable pieces yield 12 permutations of

the original pieces – a reduction by a factor of 2 which is the number of permutations of the 2 pieces with labels.

Using such argument we can solve another variation of the boxes and balls problem.

3.8.1. *Problem.* There are n balls, p_1 of them are of one kind, p_2 of them are of a second kind, . . . , and p_t of them are of a t-th kind. Any two balls of the same kind cannot be distinguished from each other while two balls of different kinds are distinguishable. In how many ways can we place the balls in n distinct boxes, no box to contain more than one ball?

Solution. Label the balls of the same kind with numbers 1, 2, 3, . . . After labelling all t kinds we have n distinguishable balls. We then have $n!$ different ways of placing these balls in the boxes, each placement corresponding to a permutation of the n different balls. Arrange these $n!$ permutations into groups in such a way that any two permutations in the same group differ from each other only by the positions of balls of the first kind. Then each group will have exactly $p_1!$ permutations since we can permute the p_1 now distinguishable balls of the first kind in $p_1!$ different ways. After removing the labels on all the balls of the first kind, the $p_1!$ permutations of each group will become indistinguishable. Therefore after the first removal the number of permutations is reduced by a factor of $p_1!$. Rearrange these $n!/p_1!$ permutations in groups in such a way that any two permutations in the same group differ from each other only by the positions of balls of the second kind. Then each group will have exactly $p_2!$ permutations which will become indistinguishable after the labels on all the balls of the second kind are taken off. Thus we have a further reduction by a factor of $p_2!$ in the number of permutations. Of the $n!/p_1!p_2!$ diffeent permutations that are left, we proceed with the removal of labels on one kind of balls after another. This will result in further reductions on the number of permutations by factors of $p_3!, p_4!,$ . . . $p_t!$. At the end when all the labels are taken off and the balls returned to their original forms, we find that there are exactly

$$\frac{n!}{p_1!\,p_2!\,\ldots\,p_t!}$$

different permutations of the original balls, each permutation corresponds

to one placement. Therefore we have

$$\frac{n!}{p_1!\,p_2!\,\ldots\,p_t!}$$

different ways to put the n given balls into n boxes. ∎

Putting this result in the language of combinatorics we have:

3.8.2. Theorem. *The number of ways to arrange n objects, where p_1 of them are of one kind, p_2 of them are of a second kind, . . . , and p_t of them are of a t-th kind ($n = p_1 + p_2 + \ldots + p_t$), is*

$$\frac{n!}{p_1!\,p_2!\,\ldots\,p_t!}$$

3.8.3. *Example.* In how many ways can the letters of the word 'tattoo' be arranged? How many of these arrangements are there in which (i) the two o's come together, (ii) the three t's come together, (iii) the two o's come together and the three t's come together, (iv) two t's come together but the third t is separated from them, (v) the three t's are all separated from each other, (vi) the three t's are all separated from each other but the two o's come together, (vii) the three t's are all separated from each other and the two o's are separated from each other?

Solution. The number of arrangements of the letters of 'tattoo' subject to no restriction is

$$6!/(3!\,2!) = 60.$$

(i) The two o's coming together may be treated as one letter (oo). The number arrangements in which the two o's come together is

$$5!/3! = 20.$$

(ii) Similarly the number of arrangements in which the three t's come together is

$$4!/2! = 12.$$

(iii) The number of arrangements in which the t's come together and the o's come together is

$$3! = 6.$$

(iv) Treating two t's together as a new letter (tt), the number of arrangements in which two t's come together is

$$5!/2! = 60.$$

Among these we have all arrangements in which all three t's come together. However these arrangements are all counted twice since the two bracketed versions t(tt) and (tt)t become the same ttt when the brackets are removed. Therefore the number of arrangements in which two t's come together while the third t is separated from them is

$$60 - 2 \times 12 = 36.$$

(v) The number of arrangements in which all t's are separated from each other is the total number of arrangements less those of (ii) and (iv). It is

$$60 - 12 - 36 = 12.$$

(vi) Applying the method used above to 'tatt(oo)' instead of to 'tattoo', we obtain the number of arrangements in which all t's are separated but the two o's come together:

$$(5!/3!) - 3! - (4! - 2 \times 3!) = 20 - 6 - 12 = 2.$$

These are obviously the arrangements 'tatoot' and 'tootat'.

(vii) The number of arrangements in which all t's are separated and all o's are separated can be obtained from the results of (v) and (vi), being $12 - 2 = 10$. ∎

3.8.4. *Example.* At an international airport, there are m counters for immigration control. Find the number of all possible schedules for n passengers to go through the counters, with each passenger going through one counter once.

Discussion. If $m = 1$ and $n = 10$, then the 10 passengers can line up in 10! ways to go through the single counter one by one. Hence with 1 counter there are 10! schedules for 10 passengers. On the other hand if $m = 10$ and $n = 1$, then the single passenger can go through any one of the ten counters that he likes. Hence with 10 counters there are 10 different schedules for 1 passenger. The matter is more complicated when both numbers m and n are greater than 1. Take, for example, $m = 3$ and $n = 5$, where the five passengers are A, B, C, D and E. Let us devise a code for every schedule. For instance, if passengers C and D go through the first counter with D following C, passengers A and E go through the second counter with E following A, and passenger B goes through the third counter, then we use the code

$$\text{CD} \mid \text{AE} \mid \text{B}$$

which is a permutation of the five letters and the two indistinguishable

separation strokes. Other schedules are similarly coded. Conversely, every such code yields a schedule, for example, the code

$$|\,BCAED\,|$$

means that no passenger goes through the first counter, all five passengers go through the second counter in the order of B, C, A, E, D and no passenger goes through the third counter. Therefore the number of schedules is the same as the number of such codes which is the number of permutations of 7 objects, two of which are alike. There are $7!/2! = 2{,}520$ different schedules for 5 passengers with 3 counters. We are now ready to solve the problem of Example 3.8.4.

Solution. Denote the passengers by capital letters A, B, ... With m counters we need $m - 1$ separation strokes |. Every schedule is now represented by a code which is a permutation of the n different letters together with the $m - 1$ separation strokes. The separation strokes divide the n letters or the code into m segments. The r-th segment is then the list of the passengers to go through the r-th counter, in the order of their appearance in the list. It may, of course, happen that some segment contains no letter at all, in which case no passenger is to go through the corresponding counter. The number of all possible codes is

$$\frac{(m+n-1)!}{(m-1)!} = (m+n-1)(m+n-2)\ldots(m)$$

which is the number of permutations of $m + n - 1$ objects (i.e. n passengers and $m - 1$ separation strokes) in which $m - 1$ of them are alike. Therefore the total number of schedules is $(m + n - 1)(m + n - 2) \ldots (m)$. ∎

3.9. CIRCULAR PERMUTATIONS

The method of putting on and taking off labels on like objects can be used effectively to solve a new type of problem.

3.9.1. *Example.* In how many ways can we paint the four seats on a merry-go-round with 4 different colours, with no two seats having the same colour.

Solution. Let us denote the 4 colours by A, B, C, D and put numbered

labels 1, 2, 3, 4 clockwise on the positions of the seats of the merry-go-round. There are 4! ways of assigning the letters to the numbers. These can be divided into groups of four each, so that the four different assignments in the same group will become indistinguishable after the removal of the numbered labels. Take one such group, for instance, with the four assignments (Fig. 3.2):

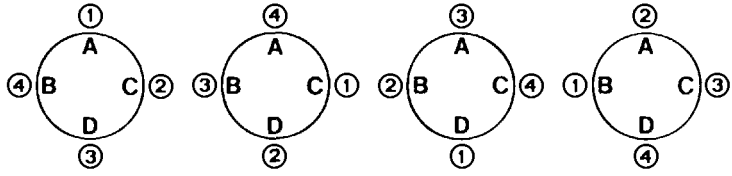


Fig. 3.2

After the removal of the labels, they all become (Fig. 3.3)

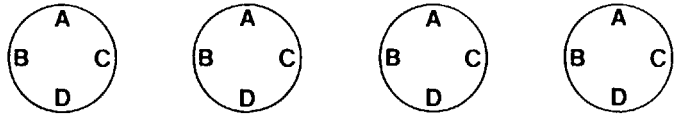


Fig. 3.3

which are just the same way of painting the seats. Therefore the reduction factor is 4 which is the number of turns of the positions to make one complete revolution. The total number of ways to paint the seats is $4!/4 = 3! = (4-1)!$ ∎

Similarly with n positions regularly distributed on the circumference of a circle and labelled clockwisely $1, 2, \ldots, n$. Turning the circle from position 1 to position 2, 2 to 3, 3 to 4, $\ldots$, $n-1$ to n and n to 1, it would take us n such turns to come back to the original position of the circle. If we follow the customary language and call an arrangement of objects on a circle a *circular permutation*, then we expect that there are $(n-1)!$ circular permutations of n distinct objects.

3.9.2. Theorem. *There are $(n-1)!$ circular permutations of n distinct objects.*

Proof. We propose to prove the theorem by induction. For $n = 1$ the theorem is obviously true. Suppose that the theorem is true for $n = r$. Given $r + 1$ distinct objects, we can arrange r of them on a circle in $(r - 1)!$ different ways. With each of these $(r - 1)!$ arrangements there are r different positions for the $(r + 1)$-th object, being positions between every two neighbouring objects of the given arrangement. By the rule of product there are $r \cdot (r - 1)! = r! = ((r + 1) - 1)!$ ways of arranging $r + 1$ objects on a circle. By the principle of induction, the theorem holds for all natural numbers $n = 1, 2, 3, \ldots$. ∎

3.10. EXERCISE

1. Consider a three-dimensional steel framework; how many different paths of length nine units are there, from one intersecting point in the framework to another, which is located two units to the right, three units back, and four units up?

2. (a) How many permutations are there of the nine letters, *A, A, A, B, B, B, C, C, C,* taken all at a time, subject to the restriction that no two *A*'s are adjacent?
   (b) What would be the answer to (a) if an additional restriction were imposed, namely that no two *B*'s are adjacent?
   What would be the answer to (b) if yet another restriction were imposed, namely that no two *C*'s are adjacent?

3. If $m > n - 2$, find the number of ways in which m positive and n negative signs can be placed in a row that no two negative signs are together.

4. (a) In how many ways can four ladies and four men be seated at a round table, if no two men are to be in adjacent seats?
   (b) Suppose the persons in (a) are four married couples. What would be the answer to the question if no husband and wife, as well as no two men, are to be in adjacent seats?

5. In how many ways can four ladies and four men sit at a round table so that each lady sits between two men?

6. In how many ways can seven beads of different colours be strung together to make a necklace?

7. How many differently coloured blocks of a fixed cubical shape can be made, if six colours are available, and a block is to have a different colour on each of its six faces?

8. In how many ways is it possible to seat eight persons at a round table, if certain two of the eight persons must not sit in adjacent seats?

3.11. COMBINATIONS

We have seen in Section 3.8 that there are $n!/r!(n-r)!$ ways of putting n balls, r of which are of one kind and $(n-r)$ of which are of another kind, in n different boxes. There are the same number of ways of putting r indistinguishable balls in n boxes with $r \leqslant n$, no box holding more than one ball. This result can be used for solving the classical problem of selecting r objects among n given objects. Similar to the solution of finding the number of subsets of a given finite set (see Example 2.5.2), we can interpret the n boxes as the n different objects. Then each placement of the r indistinguishable balls yields a selection of r objects (boxes containing a ball). Conversely every selection of r objects determines one placement of the r balls in the same manner. Therefore we have $n!/r!(n-r)!$ selections of r objects among n given objects. Each such selection is called a *combination of n objects taken r at a time*. The number $n!/r!(n-r)!$ is denoted by $C(n, r)$. Thus we have proved the following:

3.11.1. Theorem. *The number of combinations of n distinct objects taken r at a time is*

$$C(n, r) = \frac{n!}{r!\,(n-r)!}$$

3.11.2. *Example.* Three numbers are chosen among the integers 1, 2, 3, . . . , 300. How many triplets of these integers can be selected such that their sum is a multiple of 3?

Solution. The 300 integers from 1 to 300 can be divided into three sets consisting of integers which upon division by 3 give a remainder 1, 2, or 0 respectively. The three sets are:

$$A = \{1, 4, 7, 10, \ldots, 298\}$$
$$B = \{2, 5, 8, 11, \ldots, 299\}$$
$$C = \{3, 6, 9, 12, \ldots, 300\}$$

A triplet of three integers whose sum is divisible by 3 must be either of the following four types:

(i) All three integers belong to A;
(ii) All three intgers belong to B;
(iii) All three integers belong to C;
(iv) One integer from each of the sets A, B and C.

The numbers of triplets of these four types are $C(100, 3)$, $C(100, 3)$, $C(100, 3)$ and 100^3 respectively. Therefore by the rule of sum the number of all such triplets is

$$C(100, 3) + C(100, 3) + C(100, 3) + 100^3 = 1{,}485{,}100.$$ ■

3.12. EXERCISE

1. In how many ways can a group of 5 men be chosen from a squad of 9 so that two particular men will always appear?

2. How many distinct sums can be formed by adding
   (i) 3,
   (ii) 6
   of the numbers 1, 2, 4, 8, 16, 32, 64, 128, 256?

3. From 12 books, in how many ways can a selection of 5 be made,
   (i) when one specified book is always included,
   (ii) when one specified book is always excluded?

4. How many committees of 9 members can be chosen from 8 women and 6 men
   (i) without restriction,
   (ii) with exactly 1 man,
   (iii) with at least 3 men?

5. From the 26 letters of alphabet, how many different groups of 3 letters are there such that no two of the three are consecutive letters of the alphabet?

6. In how many ways is it possible to separate mn different objects into n batches with m objects in each batch?

7. Find the number of ways in which a committee of 4 can be chosen from 6 girls and 6 boys
   (a) if it must contain 2 girls and 2 boys,
   (b) if it must contain at least 1 girl and 1 boy,
   (c) if either the oldest boy or the oldest girl must be included, but not both.

8. A, B, C, D, E, F, G, H are 8 points on a plane. B, C, D are collinear but no other 3 points are collinear; E, F, G, H are concylic but no other 4 points are concylic. How many straight lines and how many circles are determined by the 8 points?

9. A, B, C, D, E are 5 points in space, of which no 3 are collinear and no 4 are coplanar. Find
   (a) the number of straight lines which join the points in pairs;
   (b) the number of planes containing 3 of the points;
   (c) the number of straight lines formed by the intersection of these planes;
   (d) the number of these planes in which any one of the given points lies.

10. A boat is to be manned by eight men, of whom two can only row on one side and one can only row on the other side; in how many ways can the crew be arranged?

11. In how many ways can four English and one French book be placed on a shelf, so that the French book is always in the middle, the selection being made from seven English and three French books?

12. Four medical tests A, B, C and D are carried out within fourteen days on a patient. A must precede B and B must precede C and D. On the days when A or B is carried out the patient must not undergo any other test, but C and D can be carried out on the same day or on different days in any order. In how many ways can the days for the tests be chosen?

13. A rectangle is divided by m lines parallel to one pair of opposite sides and n lines parallel to the other. How many rectangles are there in the figure obtained?

14. How many integers are of the form

$$a_1 a_2 a_3 \ldots a_{n-1} a_n a_{n-1} \ldots a_3 a_2 a_1$$

where $0 < a_1 < a_2 < a_3 < \ldots < a_{n-1} < a_n \leqslant 9$ and $n \geqslant 2$?

(In words: The numbers are symmetrical, having more than one digit, but none of the digits are zero and the digits are strictly increasing towards the centre. Examples: 353, 12721, 246898642)

15. Prove that the number of ways of dividing $a + b + c$ different objects into three groups containing a, b and c objects respectively is $\dfrac{(a+b+c)!}{a!\,b!\,c!}$, the three numbers a, b and c being distinct.

16. Find the coefficient of the term in $x^r y^s z^{n-r-s}$ in the expansion of $(x + y + z)^n$, where n is a positive integer.

17. A polyhedron has n faces (no two of which are congruent). Of these faces, a should be printed white, b yellow, and c green; we suppose that $a + b + c = n$. In how many different ways can this be done?

*18. If n different objects are placed round a ring, prove that the number of ways of selecting three of them, so that no selection contains two adjacent objects, is $\dfrac{n}{6}(n-4)(n-5)$.

*19. If there are n things of which four are alike and the rest are all different, find the number of combinations of these things taken $n - 3$ at a time.

*20. A number, n, of points in a plane are joined in all possible ways by straight lines, produced in both directions; no two of the lines are coincident or parallel, and no three pass through the same point. Prove that the number of points of intersection, exclusive of the n original points, is $\dfrac{1}{8}n(n-1)(n-2)(n-3)$.

*21. A man has a large supply of wooden regular tetrahedra, all of the same size. If he paints each triangular face in one of four colours, how many

differently-painted tetrahedra can be made, allowing all possible combinations of colours?
(Say that two blocks are different if they cannot be put into matching positions with identical colours on corresponding faces.)

3.13. COMBINATIONS WITH REPETITIONS

Consider yet another variation of the boxes and balls.

3.13.1. *Problem.* Let r and n be any two positive integers. There are r indistinguishable balls of the same colour, and n distinct boxes. In how many ways can we place the r balls in the n boxes if each box may hold as many balls as we wish?

Solution. Using 0 to denote a ball and 1 to denote a partition between boxes when the boxes are placed one next to the other. Every placement is then represented by a code, which is a sequence of 0's and 1's, with 0 appearing r times and 1 appearing $n - 1$ times. The $(n - 1)$ 1's of the sequence divide the sequence into n segments of 0's, the number of 0's in the p-th segment being the numbers of balls in the p-th box. For example the sequence

$$1001101 \ldots 10000$$

is the code of the placement in which no ball is put in the first box, 2 balls in the second box, none in the third, one in the fourth, . . . , and four in the n-th box. Conversely each sequence of r 0's and $(n - 1)$ 1's determines a placement of r balls in n boxes in this manner. Therefore the number of placements is the same as the number of such sequences. On the other hand each sequence is a permutation of $n + r - 1$ objects, r of which are of one kind (the 0's) and $n - 1$ of which are of another kind (the 1's). Therefore the number of placements is

$$\frac{(n + r - 1)!}{r!\,(n - 1)!} = C(n + r - 1, r). \qquad \blacksquare$$

Interpreting the n distinct boxes as n distinct kinds of objects and the number of balls in a box as the number of times in which the corresponding kinds of object is selected, we have the following classical results.

3.13.2. Theorem. *There are $C(n + r - 1, r)$ combinations of n kinds of objects taken r objects at a time.*

A sequence of p digits of 0's and q digits of 1's is called a *binary sequence of type* (p, q). The total number of digits is $p + q$ which is also called the *length* of the sequence. From the solution of Problem 3.13.1 we have the following results which are rather useful in their applications.

3.13.3. *Example.* There are $C(p + q, p) = C(p + q, q)$ binary sequences of the type (p, q).

3.13.4. *Example.* Suppose that Fig. 3.4 is the map of a city with $p + 1$ streets in the east-west direction and $q + 1$ streets in the north-south direction.

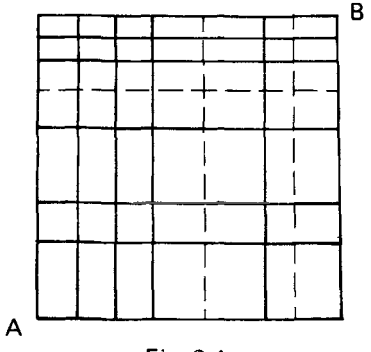


Fig. 3.4

How many different routes can a postman take from point A to point B taking the shortest distance?

Solution. Representing every east-west section between streets by 1 and every north-south section between streets by 0, the number of routes is the same as the number of binary sequences of type (p, q) which is $C(p + q, p)$.

3.13.5. *Example.* Let S be the set of all ordered triples (x_1, x_2, x_3) of

integers such that $1 \leqslant x_i \leqslant 19$, i = 1, 2, 3. Let T be the subset of S consisting of all triples such that $x_1 + x_2 + x_3 = 21$. Let U be the subset of T consisting of all triples such that $x_i < x_j + x_k$ where ijk is any permutation of 1, 2, 3. Find the number of elements of the sets S, T and U.

Solution. (a) The set S. There are 19 possibilities (the numbers 1, 2, . . . , 19) for each coordinate of triples of S. Therefore S has $19^3 = 6{,}859$ elements.

(b) The set T. Taking boxes to be coordinates and the number of ball in each box as the value of the corresponding coordinate, we see that the number of elements of T is the same as the number of different ways to put 21 balls of the same kind in 3 different boxes, each box holding one or more balls. After putting one ball in each box, we can distribute the remaining 18 balls in any way we wish. Therefore we have $C(3 + 18 - 1, 18) = C(20, 18) = 190$ placements. Hence the number of elements of T is 190.

(c) The set U. Consider the complement V of U in T. V consists of three different types of elements

(i) $(x_1, x_2, x_3) \in T$ such that $x_1 \geqslant x_2 + x_3$

(ii) $(x_1, x_2, x_3) \in T$ such that $x_2 \geqslant x_1 + x_3$

(iii) $(x_1, x_2, x_3) \in T$ such that $x_3 \geqslant x_1 + x_2$.

Let us count the elements of type (i). It follows from $x_1 + x_2 + x_3 = 21$ and $x_1 \geqslant x_2 + x_3$ that $x_1 \geqslant 11$. Using the same model as in (b) with the further restriction that the first box should hold no less than 11 balls, we find that there are $C(3 + 8 - 1, 8) = C(10, 8) = 45$ elements of type (i). Similarly there are 45 elements each of types (ii) and (iii). V has 135 elements. Hence U has $190 - 135 = 55$ elements. ∎

3.14. EXERCISE

1. Find the total number of different combinations containing 1, 2, 3, . . . , n things which can be chosen from n different things.

2. How many factors does the number $3^9 \times 5^5 \times 7^4$ have, including 1 and the number itself?

3. Ten identical coins are to be distributed among five people. Find the number of ways of distribution so that at least one person gets nothing.

4. How many unlike terms are there in the expansion of $(x + y + z)^4$?

5. Show that the number of combinations, taking $n - 2$ at a time from n things, of which three are alike and the rest different, is $\frac{n^2 - 5n + 8}{2}$.

*6. If n is a positive integer, show that the number of distinct ways in which four positive integers, of which two and only two are equal, can be chosen to have a sum $12n + 1$, is $n(18n - 7)$.

*7. Given n_1 things of one kind, n_2 of a second, n_3 of a third and n_4 of a fourth, show that, if $n_1 < r < n_2 < n_3 < n_4$, the number of groups having r things is

$$\frac{1}{6}[(r+1)(r+2)(r+3) - (r-n_1)(r-n_1+1)(r-n_1+2)].$$

*8. Consider the positive integers with five digits, namely, the integers from 10000 to 99999 inclusive. They are separated into sets. Two integers are put into the same set if their digits are the same. For example, 81332 and 31328 are in the same set while 81332 and 82132 are in different sets. How many sets are there?

3.15. BINOMIAL THEOREM

The number

$$C(n, r) = \frac{n!}{(n-r)!\,r!} = \frac{n(n-1)\ldots(n-r+1)}{1 \cdot 2 \ldots r}$$

of combinations of n distinct objects taken r at a time is also commonly known as a *binomial coefficient* and is denoted by the more convenient symbol $\binom{n}{r}$.

Thus

$$\binom{n}{r} = \frac{n(n-1)\ldots(n-r+1)}{1 \cdot 2 \ldots r}$$

They are so named because their appearance as the coefficients in the expansion of the binomial $(a+b)^n$:

$$(a+b)^1 = a+b$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$
$$= \binom{3}{0}a^3 + \binom{3}{1}a^2b + \binom{3}{2}ab^2 + \binom{3}{3}b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$
$$= \binom{4}{0}a^4 + \binom{4}{1}a^3b + \binom{4}{2}a^2b^2 + \binom{4}{3}ab^3 + \binom{4}{4}b^4$$

In general we have

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \ldots + \binom{n}{r}a^{n-r}b^r + \ldots + \binom{n}{n-1}ab^{n-1} + b^n$$

To prove the last identity we regard $(a+b)^n = (a+b)(a+b)\ldots(a+b)$ as the product of n factors each of which is the binomial $a+b$. Every term in the total expansion of this product is itself a product of n factors each of which is either a or b. Therefore it is a sequence of a's and b's of length n. The total number of such terms and hence also of such sequences is n^2. Because of the commutative law of multiplication some of these terms are equal; in fact two terms are equal if and only if they have the same number of a factors and the same number of b factors. Thus the total number of terms which are equal to $a^{n-r}b^r$ is the same as the number of sequences of $(n-r)$ a's and r b's. Therefore after collecting like terms among the n^2 terms of the expansion, the coefficient of $a^{n-r}b^r$ is $C(n,r) = \binom{n}{r}$. We have therefore proved

3.15.1. Theorem. *The binomial theorem*

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r}b^r$$

We list a number of useful identities of binomial coefficients in the following theorem:

3.15.2. Theorem. *The following identities hold:*

(a) $\binom{n}{r} = \binom{n}{n-r}$

(b) $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$ *(Pascal's rule)*

(c) $\binom{n+r+1}{r} = \binom{n+r}{r} + \binom{n+r-1}{r-1} + \ldots + \binom{n+1}{1} + \binom{n}{0}$

(d) $\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n$

(e) $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \ldots + (-1)^n \binom{n}{n} = 0$

(f) $\binom{m+n}{r} = \binom{m}{0}\binom{n}{r} + \binom{m}{1}\binom{n}{r-1} + \ldots + \binom{m}{r}\binom{n}{0}$

where $r \leqslant \min(m, n)$

(g) $\binom{n}{0}^2 + \binom{n}{1}^2 + \ldots + \binom{n}{n}^2 = \binom{2n}{n}$

(h) $\binom{n}{1} + 2\binom{n}{2} + \ldots + n\binom{n}{n} = n \times 2^{n-1}$.

The proof of these identities are left as an exercise. The identity (b), known as *Pascal's rule*, is illustrated by the so called *Pascal's triangle* below.

$$
\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
\ldots\ldots\ldots\ldots\ldots \\
1 \ldots \binom{n}{r-1}\binom{n}{r} \ldots 1 \\
1 \ldots\ldots \binom{n+1}{r} \ldots\ldots 1 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{array}
$$

This configuration of numbers is named after the French philosopher, Blaise Pascal (1623 – 1662), but is also known to a number of mathematicians before his time, for example, the Arab Al-Kashi (15th century) and the Chinese 朱世傑 (14th century) 楊 輝 (13th century) and 賈憲 (11th century).

3.15.3. *Example.* Determine which rows of the Pascal's triangle have the

property that there are three consecutive binomial coefficients in an arithmetic progression.

Solution. Upon examination of the first few rows of the Pascal's triangle, we shall discover that we have the first two such triples on the 7th row, namely

$$\binom{7}{1} = 7 \ , \binom{7}{2} = 21 \ , \binom{7}{3} = 35$$
$$\binom{7}{4} = 35 \ , \binom{7}{5} = 21 \ , \binom{7}{6} = 7.$$

Now 7 is an odd number. This may lead us to think that such triples can be found on all odd numbered rows. Going through the 10 coefficients on the 9th row, we realize that we have made a wrong guess. The next guess, that we shall find such triples on the n-th row where n is prime, will also turn out to be erroneous. Instead of making other wild guesses, let us try to find a necessary condition for the existence of such triples.

Suppose that for a fixed $n \geqslant 7$ there are three consecutive binomial coefficients

$$\binom{n}{r-1}, \binom{n}{r}, \binom{n}{r+1} \quad (r < n)$$

on the n-th row of the Pascal's triangle, which are in an arithmetic progression. Then

$$2\binom{n}{r} = \binom{n}{r-1} + \binom{n}{r+1}$$

$$\therefore \quad \frac{2 \times n!}{(n-r)!\, r!} = \frac{n!\,[(r+1)r + (n-r+1)\,(n-r)]}{(n-r+1)!\,(r+1)!}$$

Hence $2(n-r+1)\,(r+1) = (r+1)r + (n-r+1)\,(n-r)$

$\therefore \quad n^2 - 4nr + 4r^2 = n + 2$

$\therefore \quad n + 2 = (n-2r)^2.$

Thus a necessary condition for the existence of such triples on the n-th row is that $n + 2$ is a perfect square. In other words, there are no such triples except on the 7th ($n + 2 = 3^2$), the 14th ($n + 2 = 4^2$), the 23rd ($n + 2 = 5^2$), . . . ($n + 2 = p^2$), . . . rows of the triangle. Fortunately this condition turns out to be sufficient. To see this, let $n + 2 = p^2$ where p is an integer not less than 3. Then the integers p and n must have the same parity, i.e. they are at the same time odd or even. It follows that

$$r = \frac{1}{2}(n - p)$$

is a positive integer less than $n/2$ and

$$n + 2 = (n - 2r)^2 = n^2 - 4nr + 4r^2.$$

Retracing the previous steps backward, we obtain

$$2\binom{n}{r} = \binom{n}{r-1} + \binom{n}{r+1}$$

Therefore the triple

$$\binom{n}{r-1}, \binom{n}{r}, \binom{n}{r+1}$$

and by symmetry also the triple

$$\binom{n}{n-r-1}, \binom{n}{n-r}, \binom{n}{n-r+1}$$

are in an arithmetic progression. The final solution to the problem is therefore that there are always two triples of consecutive binomial coefficients which are in an arithmetic progression on each $(p^2 - 2)$-th row of the triangle where $p = 3, 4, 5, \ldots$ and there are no such triples elsewhere in the triangle. ■

3.15.4. *Example.* Find the value of the expression

$$f(x) = \frac{(1 - x + 3x^2)^8}{(1 - 5x)^4}$$

for values of x which are so small that the power of x above the square may be neglected.

Solution. It is required of us to approximate $f(x)$ when x is small enough for us to neglect all terms involving $x^3, x^4, \ldots$. Let us first approximate the numerator and the denominator separately. For the numerator we have

$$(1 - x + 3x^2)^8 = [1 - x(1 - 3x)]^8$$
$$= 1 - 8x(1 - 3x) + 28x^2(1 - 3x)^2 + \text{terms of higher orders}$$
$$= 1 - 8x + 52x^2 + \text{terms of higher orders.}$$

For the denominator, let us first consider the infinite geometric progression

$$1 + 5x + (5x)^2 + (5x)^3 + (5x)^4 + \ldots$$

For small values of x we have

$$\frac{1}{1 - 5x} = 1 + 5x + (5x)^2 + \text{terms of higher orders.}$$

Therefore for small values of x,

$$\frac{1}{(1-5x)^4} = (1 + 5x + (5x)^2)^4 + \text{terms of higher orders}$$
$$= 1 + 20x(1 + 5x) + 150x^2(1 + 5x)^2 + \text{terms of higher orders}$$
$$= 1 + 20x + 250x^2 + \text{terms of higher orders.}$$

Therefore for small values of x,

$$f(x) = (1 - 8x + 52x^2 + \text{terms of higher orders})(1 + 20x + 250x^2 + \text{terms of higher orders})$$
$$= 1 - 12x + 142x^2 + \text{terms of higher orders.}$$

After dropping all terms of higher orders, we conclude that for small values of x

$$f(x) = \frac{(1 - x + 3x^2)^8}{(1 - 5x)^4}$$

is practically the same as $1 - 12x + 142x^2$. ∎

3.16. EXERCISE

1. If n and r are natural numbers such that $1 \leqslant r < n$, prove that
$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}.$$
Deduce that $\binom{n}{r}$ is an integer.

2. Prove that $\sum_{r=0}^{k} (-1)^r \binom{n}{r} = (-1)^k \binom{n-1}{k}$ for any positive integers n and k such that $n > k$.

3. Prove that $\sum_{r=0}^{n-1} \binom{n}{r}\binom{n}{r+1} = \frac{(2n)!}{(n+1)!\,(n-1)!}$ for any positive integer n.
(Hint: Consider the product $(1+x)^n\,(1+\frac{1}{x})^n$.)

4. Prove that $\sum_{k=0}^{n} \binom{r+k}{r} = \binom{r+n+1}{r+1}$ for any positive integers r and n.

5. Prove that, for any positive integer n,

$$\prod_{r=1}^{n} \binom{2r}{2} = n! \prod_{r=1}^{n} (2r-1) = \frac{(2n)!}{2^n}.$$

6. (a) Find the sum $\sum_{i=1}^{n} i$.

   (b) Prove that $2^n \geqslant \binom{n}{0} + \binom{n}{1}$ for any positive integer n.

   (c) Using (a) and (b), or otherwise, prove that $2^{n(n+1)/2} \geqslant (n+1)!$ for any positive integer n.

7. Let $f(r) = \sum_{i=0}^{n-r} \binom{n}{i}\binom{n}{r+i}$ where n is a positive integer.

   (a) By considering the expansion of $(1+x)^{2n}$, or otherwise, show that

$$f(r) = \frac{(2n)!}{(n+r)!\,(n-r)!}$$

   (b) By considering the expansion of $(1+x)^{3n}$ and using (a), or otherwise, show that

$$\sum_{r=0}^{n} \binom{n}{r} f(r) = \frac{(3n)!}{n!\,(2n)!}$$

8. Prove that $\sum_{r=0}^{m} \binom{n-m}{r}\binom{n+m}{m-r} = \binom{2n}{m}$ for any natural numbers n and m such that $n > m$.

*9. If n is a positive integer, prove that

$$\sum_{r=0}^{n} (-1)^r \binom{n}{r}^2 = \begin{cases} 0 \text{ if } n \text{ is odd,} \\ \dfrac{(-1)^m\,(2m)!}{(m!)^2} & \text{if } n\ (=2m) \text{ is even.} \end{cases}$$

*10. Prove that $\sum_{i=0}^{r} (-1)^i \binom{n}{i}\binom{n}{r-i} = \begin{cases} 0 \text{ if } r \text{ is odd,} \\ (-1)^{r/2}\binom{n}{\frac{r}{2}} \text{ if } r \text{ is even,} \end{cases}$

where r and n are positive integers such that $r \leqslant n$.

*11. If n is a positive integer, prove that

(a) $\sum_{r=0}^{n} 2^r \binom{n}{r} = 3^n$ and

(b) $\sum_{r=0}^{n} (-2)^r \binom{n}{r} = (-1)^n$.

*12. Let m and n be positive integers such that $n > 1$.

(a) Prove that $\left(\frac{n^2}{n^2-1}\right)^n > 1 + \frac{1}{n}$

and deduce that $\left(1 - \frac{1}{n}\right)^{-n} > \left(1 - \frac{1}{n+1}\right)^{-n-1}$

(b) Prove that $\left(1 + \frac{1}{m}\right)^m < \left(1 + \frac{1}{m+1}\right)^{m+1}$

(c) Prove that $\left(1 + \frac{1}{n}\right)^n < \left(1 - \frac{1}{n}\right)^{-n}$

and deduce that $\left(1 + \frac{1}{m}\right)^m < \left(1 - \frac{1}{n}\right)^{-n}$.

13. Write down the following expansions:

(a) $\left(x + \frac{1}{x}\right)^4$.

(b) $(2x - 5)^5$.

14. Find the term independent of x in the expansion of $\left(x^3 - \frac{1}{x^2}\right)^5$.

15. The coefficient of x^6 in the expansion of $(1 + kx)^{10}$ in powers of x is equal to the coefficient of x^8. Find the possible values of k.

16. If three consecutive coefficients in the expansion of $(1 + x)^n$, n being a positive integer, are $2a$, $15a$ and $70a$, find the values of n and a.

17. Expand $(1 + 2x + 3x^2)^4$ in ascending powers of x, calculating the coefficients as far as the term in x^3.

18. In the expansion of $(1 + \lambda x - 3x^2)^6$, the coefficients of the terms involving x^2 and x^3 are 42 and 20 respectively. Determine the value of λ and the coefficient of the term involving x^4.

19. Find the coefficient of x in the expansion of $(1+\frac{1}{x})^2(2-3x)^6$.

20. Show that when $(1-2x)^{18}(1+3x)^{17}$ is expanded, the coefficient of the term involving x^2 is zero.

21. You are given that, when $(x+a)^3(x-b)^6$ is expanded in powers of x, the coefficient of x^8 is zero and the coefficient of x^7 is -9. Find all the possible values of a and b.

22. Find the values of the positive integers m and n, so that the expansions of $(1-x^2)^{15}$ and of $(1+2x)^m(1-3x)^n$ may have the same coefficients of x and of x^2.

23. (a) Prove the identity
$$x^5+y^5 \equiv (x+y)^5 - 5xy(x+y)^3 + 5x^2y^2(x+y).$$
(b) A positive number x and a negative number y exist such that
$$x+y=1 \quad \text{and} \quad x^5+y^5=2101.$$
(i) Use the identity in (a) to form a quadratic equation in xy.
(ii) Deduce that $xy=-20$.
(iii) Hence find x and y.

*24. It is known that x, a and n are positive integers and that the first three terms in the binomial expansion of $(x+a)^n$ in descending powers of x are 729, 2916 and 4860 respectively. Find the values of x, a and n.

*25. If y denotes $x+\frac{1}{x}$, express $x^7+\frac{1}{x^7}$ in the form
$$y^7+ay^5+by^3+cy,$$
where a, b and c are numerical coefficients.
(Hint: Consider $(x+\frac{1}{x})^7$.)

*26. If z denotes $x-\frac{1}{x}$, prove that
$$(x^7-\frac{1}{x^7}) \div (x-\frac{1}{x})$$
can be expressed in the form $z^6+az^4+bz^2+c$, and find the values of a, b and c.

27. (a) Write down the first four terms in the expansion of

$$\frac{3}{(1-x)(1+2x)}$$

in ascending powers of x.

(b) State the range of values of x for which the expansion is valid.

4. *Arithmetic*

In this chapter we shall put together in a systematic and rigorous manner some results of elementary arithmetic that are familiar to our readers.

4.1. ABSOLUTE VALUE

We shall be dealing exclusively with integers. These are whole numbers

$$0, \pm 1, \pm 2, \pm 3, \ldots$$

The set of all integers is customarily denoted by the bold-faced letter $\mathbb{Z}$. The *absolute value* of an integer n is given as the natural number

$$|n| = \begin{cases} n & \text{if } n \geqslant 0 \\ -n & \text{if } n \leqslant 0. \end{cases}$$

Thus $|3| = 3$, $|-7| = 7$. Another way to define the absolute value is to put

$$|n| = \sqrt{n^2}.$$

Some useful properties of absolute value are summarized in the following theorem.

4.1.1. Theorem. *Let a and b be integers. Then the following statements hold:*

(a) $-|a| \leqslant a \leqslant |a|$

(b) $|ab| = |a||b|$

(c) $|a + b| \leqslant |a| + |b|$ *(triangle inequality)*

(d) $|a - b| \geqslant |a| - |b|$.

Proof. (a) Given an integer a, it is either $0 \leqslant a$ or $a < 0$. In the former case, $-|a| \leqslant 0 \leqslant a = |a|$. In the latter case $-|a| = a < 0 < |a|$. Therefore the statement holds in both cases.

(b) This follows from

$$|ab| = \sqrt{(ab)^2} = \sqrt{a^2}\,\sqrt{b^2} = |a||b|$$

(c) This very useful inequality follows from

$$\begin{aligned} |a + b|^2 &= (a + b)^2 = a^2 + 2ab + b^2 \\ &= |a|^2 + 2ab + |b|^2 \end{aligned}$$

$$\leqslant |a|^2 + 2|ab| + |b|^2$$
$$= |a|^2 + 2|a||b| + |b|^2$$
$$= (|a| + |b|)^2$$

(d) Since $|a| = |b + (a - b)| \leqslant |b| + |a - b|$, the inequality follows. ∎

4.2. EXERCISE

1. Prove that $|a| < b$ if and only if $-b < a < b$.

2. Find the integral solutions of the inequality $|x + 4| < 2$.

3. Solve the inequality $\left|\frac{3x + 4}{x - 7}\right| < 1$ for integral values of x.

4. If $|x - a| + |x - b| < c$, prove that $|b - a| < c$.

5. If $|x - a| < c$ and $|y - b| < c$, show that
   (i) $|(x + y) - (a + b)| < 2c$,
   (ii) $|(x - y) - (a - b)| < 2c$.

4.3. DIVISIBILITY

We recall some familiar definitions and terminology. An integer b is said to be *divisible* by a non-zero integer a if there is some integer c such that $b = ac$. In this case we write $a|b$. For $a|b$ we sometimes also say that a is a *divisor* of b, a is a *factor* of b or b is a *multiple* of a. If b is not divisible by a then we write $a \nmid b$. When we write $a|b$ or $a \nmid b$ we shall always tacitly assume that $a \neq 0$. Thus $3|12$ and $3 \nmid 16$.

The following theorem puts together some of the most familiar facts about divisibility.

4.3.1. Theorem. *For integers a, b, c and d the following statements hold:*
(a) $a|0, \quad \pm 1|a, \quad a|a, \quad a|-a$
(b) $a|1$ if and only if $a = \pm 1$
(c) If $a|b$ and $b|c$, then $a|c$
(d) If $a|b$ and $c|d$, then $ac|bd$
(e) $a|b$ and $b|a$ if and only if $a = \pm b$

(f) If $a \mid b$ and $b \neq 0$, then $|a| \leqslant |b|$

(g) If $a \mid b$ and $a \mid c$, then $a \mid (bx + cy)$ for any integers x and y.

Proof. We leave the proof of (a), (c), (d), (g) as an exercise.

(b) If $a = \pm 1$, then obviously $a \mid 1$. Conversely, let $a \mid 1$. Then $1 = ax$ for some integer x. This means that the integer a has a reciprocal x which is also an integer. Since ± 1 are the only integers that have this property, $a = \pm 1$.

(e) Suppose that $a \mid b$ and $b \mid a$. Then we can find integers p and q such that $b = pa$ and $a = qb$. Hence $a = qb = q(pa) = (pq)a$. It follows that $pq = 1$ and hence $q = \pm 1$. Therefore $a = \pm b$. The converse is obvious.

(f) Let $a \mid b$ and $b \neq 0$. Then $b = ax$ for some integer x. Since $b \neq 0$, we must conclude that $x \neq 0$. Taking absolute value on both sides we get $|b| = |a||x|$. Since $|x| \geqslant 1$, we get $|a| \leqslant |b|$. ∎

Let us draw some immediate consequences from the theorem. It follows from (a) that the integer 0 has an infinite number of divisors. On the other hand by (b), the integers 1 and -1 have the least number of divisors, namely only 1 and -1. By (f) a non-zero integer b has only a finite number of divisors, among them the trivial ones being 1, -1, b and $-b$. Statements (d) and (g) can be generalized:

(d') If $a_i \mid b_i$ for $i = 1, 2, \ldots, n$, then $a_1 a_2 \ldots a_n \mid b_1 b_2 \ldots b_n$

(g') If $a \mid b_i$ for $i = 1, 2, \ldots, n$, then $a \mid (b_1 x_1 + b_2 x_2 + \ldots + b_n x_n)$ for any n integers $x_1, x_2, \ldots, x_n$.

The notions of common divisor and greatest common divisor are the next familiar terms that we wish to recapitulate. Let a and b be integers. A *common divisor* of a and b is an integer which is a divisor of both a and b. If a and b are not both zero, then they can only have a finite number of common divisors. Therefore among the common divisors of a and b there is one which is the largest. We call this positive integer the *greatest common divisor* of a and b and denote it by the abbreviation $\gcd(a, b)$. Take for example the integers 12 and -40. The divisors of 12 are ± 1, ± 2, ± 3, ± 4, ± 6, ± 12. The divisors of -40 are ± 1, ± 2, ± 4, ± 5, ± 8, ± 10, ± 20, ± 40. The common divisors of 12 and -40 are therefore ± 1, ± 2, ± 4 and the greatest among them is 4. Therefore $\gcd(12, -40) = 4$.

The gcd 4 is divisible by all common divisors ± 1, ± 2, ± 4 of the numbers 12 and -40, a fact which can be seen also in the light of Theorem 4.3.1.

(g) and the equation

$$4 = (12) \times (-3) + (-40) \times (-1).$$

Similarly from

$$\gcd(60, 42) = 6 = (60) \times (-2) + (42) \times (3),$$

we see that the gcd of 60 and 42 is divisible by all common divisors of 60 and 42. We shall prove this particular property of gcd in a later section.

4.4. EUCLIDEAN ALGORITHM

We know that when one positive integer b is divided by another positive integer a, it either leaves zero or a positive residue r, which is less than a, as the remainder of the division. This is illustrated by Fig. 4.1 below:

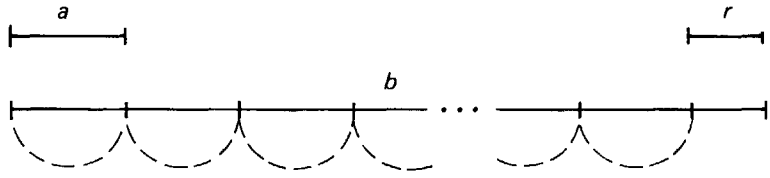


Fig. 4.1

By laying off repeatedly segments of length a from the segment b, at the end we shall arrive at having either nothing left or a residue segment of length r which is shorter than the segment a. If it takes laying off q segments, then

$$r = b - a - a - \ldots - a = b - qa \quad \text{where } 0 \leqslant r < a.$$

Since repeated subtraction is the same as division, we may call q the *quotient* and r the *remainder* of the division of b by a.

We call the above process of calculation the *Euclidean algorithm*, after the Greek mathematician Euclid (*circa* 350 B.C.), who was the first to realize its fundamental importance and used it to lay the foundation for his theory of numbers. The word algorithm, which now means a rule of operation, is a derivation of the name of an Arab mathematician al'Khowarizmi (9th century). From the title of his book, *Al-jabr wa'l muqābalah*, has also come the word algebra, for it is from this book that Europe later learned the branch of mathematics bearing this name.

Let us now formulate the above discussion into a theorem and prove it by means of the well-ordering principle (see Section 2.2).

4.4.1. Theorem. *Euclidean algorithm. If a and b are integers where $a \neq 0$, then there are unique integers q and r such that*

$$b = qa + r, \quad 0 \leqslant r < |a|.$$

Discussion. The theorem has two parts. Firstly it stipulates the existence of the integers q and r with the required property. Secondly it claims uniqueness of q and r which means that there is exactly one such pair of integers. Therefore the proof of the theorem must be given in two parts, the existence part and the uniqueness part, though it is immaterial which part comes first.

Proof. We consider separately the two cases where $a > 0$ and where $a < 0$.

(a) Existence ($a > 0$). Consider the set S of all natural numbers of the form $b - xa$ where $x \in \mathbb{Z}$,

i.e. $$S = \{b - xa \mid x \in \mathbb{Z} \text{ and } b - xa \geqslant 0\}.$$

In order to be able to apply the well-ordering principle to S, we have to show that S is non-empty. Take $x = -|b|$. Then

$$b - xa = b + |b|a \geqslant b + |b| \geqslant 0$$

since $a \geqslant 1$. Therefore $S \neq \phi$. By the well-ordering principle, S has a least element $r \in \mathbb{N}$. Obviously

$$r \geqslant 0 \text{ and } r = b - qa \text{ for some } q \in \mathbb{Z}.$$

We want to show that the pair q and r have the required property. We have seen already that

$$b = qa + r \text{ and } 0 \leqslant r,$$

it remains only to show that $r < |a| = a$. Suppose to the contrary that $r \geqslant a$. Then

$$s = b - (q + 1)a = (b - qa) - a = r - a \geqslant 0$$

would be an element of S. On the other hand

$$s = r - a < r$$

which contradicts the fact that r is the least element of S. Therefore we conclude that $0 \leqslant r < a$. This completes the proof of existence of q and r with the required property.

(b) Uniqueness ($a > 0$). Suppose that there are two pairs q, r and q', r' of integers with the required property,

$$b = qa + r \qquad 0 \leqslant r < |a| = a$$
$$b = q'a + r' \qquad 0 \leqslant r' < |a| = a$$

Then it follows from $r - r' = (q' - q)a$, $0 \leqslant r < a$ and $-a < -r' \leqslant 0$

that $-a < r - r' < a$

$\therefore \quad -a < (q' - q)\, a < a$

$\therefore \quad -1 < q' - q < 1.$

Since q and q' are integers we must have

$$q' - q = 0$$

and hence also

$$r - r' = (q' - q)\, a = 0.$$

Therefore $q = q'$ and $r = r'$. This proves that there is just one such pair of integers q and r. By (a) and (b) the theorem holds for $a > 0$.

For the case where $a < 0$, we apply what we have shown to $|a|$ and b. For $|a|$ and b there are unique integers p and r such that

$$b = p|a| + r, \quad 0 \leqslant r < |a|.$$

By putting $q = -p$, we have unique q and r such that

$$b = qa + r, \quad 0 \leqslant r < |a|.$$

Our proof is now complete. ∎

4.5. EXERCISE

1. If $a_1 a_2 | b$, show that $a_1 | b$ and $a_2 | b$.
   Is the converse of this result true? If so, prove it; if not, produce a counter-example.

2. Find the positive integer which has 8 positive divisors, such that the product of the divisors is 331776.

3. How many distinct positive divisors does the number 18000 have (including 1 and 18000)?

4. Use the Euclidean algorithm to establish that
   (a) every odd integer is either of the form $4n + 1$ or $4n + 3$;
   (b) the square of any integer is either of the form $3n$ or $3n + 1$;
   (c) the cube of any integer is of the form $9n$, $9n + 1$ or $9n + 8$, where n is an integer.

*5. Prove that (i) $\binom{n}{r}$ is odd if $n = 2^m - 1$

and (ii) $\binom{n}{r}$ is even if $n = 2^m$

where m, n and r are positive integers with $r < n$.

6. Find the quotient q and the remainder r for the following values of a and b: (i) $a = 2, b = 1$;
   (ii) $a = -3, b = 4$;
   (iii) $a = 5, b = -6$;
   (iv) $a = -4, b = -1$.

7. Prove that no integer in the sequence 11, 111, 1111, . . . is a perfect square.
   (Hint: A typical term 111 . . . 111 can be written as 111 . . . $108 + 3 = 4n + 3$ for some integer n.)

8. Show that any integer of the form $6m + 5$ is also of the form $3n + 2$, where m and n are integers, but not conversely.

9. If an integer is simultaneously a square and a cube (as is the case with $729 = 27^2 = 9^3$), prove that it must be either of the form $7n$ or $7n + 1$, where n is an integer.

10. If a and b are integers with $b > 0$, prove that there exist unique integers q and r satisfying $a = qb + r$, where $2b \leqslant r < 3b$.

11. If n is a positive integer, show that every positive integer m can be expressed uniquely in the form $m = \Sigma_{i=0}^{N} C_i n^i$, where the C_i's are natural numbers not exceeding $n - 1$.

12. For any positive integer n, prove that $\dfrac{n(n+1)(2n+1)}{6}$ is an integer.
    (Hint: By the Euclidean algorithm, n has one of the forms $6m$, $6m + 1$, . . . , $6m + 5$; establish the result in each of these six cases.)

4.6. THE GREATEST COMMON DIVISOR

We shall always assume that the integers a and b are not both zero whenever we speak of their greatest common divisor. We recall that the greatest common divisor $\gcd(a, b)$ of two integers a and b is the largest among all the common divisors of a and b. It is also known (see Section 4.3) that given a and b, this positive integer exists and is unique.

4.6.1. Theorem. *Let a and b be integers. Then* $\gcd(a, b)$ *is the least positive integer of the form $ax + by$ where x, y are integers.*

Proof. Consider the sets S of all positive integers of the form $ax + by$, i.e.

$$S = \{ax + by \mid x, y \in \mathbb{Z} \text{ and } ax + by > 0\}$$

Since a and b are not both zero, $|a| + |b|$ is a positive integer and belongs to S. By the well-ordering principle the non-empty set S has a least element d which, by definition, is the least positive integer of the form $ax + by$. It remains to show that $d = \gcd(a, b)$; in other words, to show (i) that d is a common divisor of a and b and (ii) that if c is a positive common divisor of a and b, then $c \leqslant d$.

(i) Since $d \in S$, we can write $d = ax + by$ for some integers x, y. Applying Euclidean algorithm to d and a, we get $a = qd + r$ where either $r = 0$ or $0 < r < d$. Of the two possibilties for the remainder r, the latter $0 < r < d$, turns out to be impossible for $r = a - qd = a - q(ax + by) = a(1 - qx) + b(-qy)$ would belong to S and would be strictly less than the least element d of S. Therefore we must conclude that $r = 0$. Hence $a = qd$ and $d \mid a$. Similarly we can prove that $d \mid b$. Therefore d is a common divisor of a and b.

(ii) Let c be a positive common divisor of a and b. Then it follows from Theorem 4.3.1.(g) that $c \mid (ax + by)$. Therefore $c \mid d$. By statement (f) of the same theorem we get $c = |c| \leqslant |d| = d$.

(i) and (ii) together complete the proof. ∎

4.6.2. Theorem. *Given integers a and b. A positive integer d is the greatest common divisor of a and b if and only if the following two conditions are satisfied:*

(i) $d \mid a$ and $d \mid b$ (i.e. d is a common divisor),

(ii) If $c \mid a$ and $c \mid b$, then $c \mid d$ (i.e. d is divisible by all common divisors).

Proof. (a) We first prove that if $\gcd(a, b) = d$, then d satisfies conditions (i) and (ii). By definition (i) is satisfied. By Theorem 4.6.1, $d = ax + by$ for some integers x and y. If $c \mid a$ and $c \mid b$, then $a = cs$ and $b = ct$ for some integers s and t. Therefore $d = ax + by = csx + cty = c(sx + ty)$. Hence $c \mid d$ and condition (ii) is satisfied.

(b) We prove conversely that if d satisfies conditions (i) and (ii), then $d = \gcd(a, b)$. By (i), d is a common divisor of a and b. It remains to show it is the largest among all positive common divisors. Suppose $c \mid a$, $c \mid b$ and $c > 0$. Then by (ii), $c \mid d$. Hence $c = |c| \leqslant |d| = d$

The proof of the theorem is now complete. ∎

Of particular interest is the case where $\gcd(a, b) = 1$. In such case, we say that a and b are *relatively prime*. For example, 2 and 5 are relatively prime, so are 36 and -55.

4.6.3. *Example.* Prove that integers a and b are relatively prime if and only if there are integers x and y such that $ax + by = 1$.

Proof. We shall use Theorem 4.6.1 and the fact that 1 is the least positive integer. Suppose a and b are relatively prime, then $\gcd(a, b) = 1$. By Theorem 4.6.1, $1 = ax + by$ for some integers x and y. Conversely if $1 = ax + by$ for some $x, y \in \mathbb{Z}$, then the number 1 is the least positive integer of the form $as + bt$ for all possible $s, t \in \mathbb{Z}$. By Theorem 4.6.1, $1 = \gcd(a, b)$. ∎

In general $a|bc$ does not imply $a|b$ or $a|c$. Take for example, $a = 4$, $b = 6$, $c = 10$. Then $6 \times 10 = 4 \times 15$; hence $4|(6 \times 10)$ but $4 \nmid 6$ and $4 \nmid 10$. On the other hand, we have $4|(5 \times 12)$, $\gcd(4, 5) = 1$ and $4|12$. This leads us to the following example.

4.6.4. *Example.* Show that if $a|bc$ and $\gcd(a, b) = 1$, then $a|c$.

Proof. It follows from the assumption $\gcd(a, b) = 1$ that $1 = ax + by$ for some $x, y \in \mathbb{Z}$. Then $c = a(cx) + (bc)y$. Since $a|a$ and $a|bc$, it follows that $a|c$. ∎

4.7. EXERCISE

1. Prove that $4 \nmid (n^2 + 2)$ for any integer n.

2. If $a|b$, show that $(-a)|b$, $a|(-b)$ and $(-a)|(-b)$.

3. Given integers, a, b and c, prove that $a|b$ if and only if $ac|bc$, where $c \neq 0$.

4. Prove or disprove that if $a|(b + c)$, then either $a|b$ or $a|c$.

5. Establish that $2|n(n+1)$ while $3|n(n+1)(n+2)$ for an arbitrary integer n.

6. If n ($\neq 1$) is an integer such that $2 \nmid n$ and $3 \nmid n$, show that $24|(n^2 - 1)$.

7. If k, m and n are positive integers, show that gcd $(m + kn, n)$ = gcd (m, n).

8. For any positive integer n and any integer m, prove that gcd $(m, m+n)$ divides n.

9. If x and y_i are relatively prime for all $i = 1, 2, \ldots, n$, prove that x and $y_1 y_2 \ldots y_n$ are relatively prime.

10. If m and n are integers, not both of which are zero, prove that
gcd $(-m, -n)$ = gcd $(-m, n)$ = gcd $(m, -n)$ = gcd (m, n).

11. For a non-zero integer n, show that
(i) gcd $(n, 0) = |n|$,
(ii) gcd $(n, n) = |n|$,
(iii) gcd $(n, 1) = 1$.

12. For any integer n, prove that one of the integers n, $n + 2$, $n + 4$ is divisible by 3.
(Hint: By the Euclidean algorithm the integer n must be of the form $3m$, $3m + 1$ or $3m + 2$.)

13. Given integers m and n, prove that gcd$(m, n) = am + bn$ where gcd$(a, b) = 1$.

14. Given integers m, n and c, prove that there exist integers a and b for which $c = am + bn$ if and only if gcd $(m, n)|c$.

*15. (a) If m and n are positive integers with $m > n$, prove that
gcd $(m + n, m - n)$ = gcd (m, n) or 2 x gcd (m, n)
(Hint: Let d_1 = gcd (m, n) and d_2 = gcd $(m + n, m - n)$. Show that $d_1 | d_2$ and $d_2 | 2d_1$.)
(b) Determine when the right-hand side of the equality in (a) is gcd (m, n) and when 2 x gcd (m, n).

*16. If a and b are given integers, not both zero, prove that the set

$$S = \{ax + by : x \text{ and } y \text{ are integers}\}$$

is precisely the set of all multiples of gcd (a, b).

4.8. THE LEAST COMMON MULTIPLE

The parallel concept of the greatest common divisor is known as the least common multiple. Given two non-zero integers a and b, the *least common multiple* of a and b is the smallest positive common multiple of a and b and is denoted by $\text{lcm}(a, b)$. Between the two parallel concepts we have the following theorems.

4.8.1 Theorem. *For positive integers a and b*, $\gcd(a, b)\ \text{lcm}(a, b) = ab$.

Proof. Let $d = \gcd(a, b)$. Then

$$d = ax + by,\ a = dr \text{ and } b = ds$$

for some integers r, s, x, y. Then

$$m = ab/d = as = br$$

is a positive common multiple of a and b. It remains for us to prove that $m = \text{lcm}(a, b)$. Suppose c is a positive multiple of a and b. Then

$$c = au = bv$$

for some u, v $\in \mathbb{Z}$. Dividing c by m we have

$$\frac{c}{m} = \frac{cd}{ab} = \frac{c(ax + by)}{ab} = \frac{c}{b}x + \frac{c}{a}y$$

$$= vx + uy.$$

This means that c/m is an integer; hence $m \mid c$. Therefore $m \leqslant c$ ∎

4.9. AN EFFECTIVE DIVISION ALGORITHM FOR THE EVALUATION OF GCD

The crudest method of evaluating the gcd is a search among all possible divisors. This is obviously very tedious. A very effective method is given by Euclid in his famous book *Elements*. The method consists of a finite number of steps given as follows:

Let a and b be two integers. Since $\gcd(a, 0) = \gcd(a, a) = |a|$ and $\gcd(a, b) = \gcd(b, a) = \gcd(|a|, |b|)$, we may assume that $0 < a < b$. Then

we proceed.

Step 1. Apply Euclidean algorithm to get

$$b = q_1 a + r_1 , \quad 0 \leqslant r_1 < a.$$

If $r_1 = 0$, then we stop and get $\gcd(a, b) = a$. Otherwise we proceed to the next step.

Step 2. Divide a by r_1 to get

$$a = q_2 r_1 + r_2 , \quad 0 \leqslant r_2 < r_1 .$$

If $r_2 = 0$, then we stop and get $\gcd(a, b) = r_1$. Otherwise we proceed to the next step.

Step 3. Divide r_1 by r_2 to get

$$r_1 = q_3 r_2 + r_3 , \quad 0 \leqslant r_3 < r_2 .$$

If $r_3 = 0$, then we stop and get $\gcd(a, b) = r_2$. Otherwise we proceed to the next step.

Step 4. . . . . . .

. . . . . .

Step k. Divide r_{k-2} by r_{k-1} to get

$$r_{k-2} = q_k r_{k-1} + r_k , \quad 0 \leqslant r_k < r_{k-1} .$$

If $r_k = 0$ then we stop and get $\gcd(a, b) = r_{k-1}$. Otherwise we proceed to the next step.

Step k + 1. . . . . . .

. . . . . .

Let us apply the above procedure to a pair of positive integers, say $a = 612$, $b = 4138$.

Step 1. $4138 = 6 \times 612 + 466$
Step 2. $612 = 1 \times 466 + 146$
Step 3. $466 = 3 \times 146 + 28$
Step 4. $146 = 5 \times 28 + 6$
Step 5. $28 = 4 \times 6 + 4$
Step 6. $6 = 1 \times 4 + 2$
Step 7. $4 = 2 \times 2 + 0$

In this particular case the procedure terminates at step 7 and we obtain $\gcd(612, 4138) = 2$.

Let us now examine the division algorithm in its general form. According to the description, if the procedure stops at step n with a vanishing

remainder r_n, then the last non-zero remainder r_{n-1} will be the gcd of the given numbers a and b. But can we be sure that the procedure will ever stop? Supposing that it stops at step n, is r_{n-1} really the greatest common divisor of a and b? In order to justify the proposed procedure we shall have to answer the following questions:

(1) Does the division algorithm terminate after a finite number of steps?

(2) If r_n is the first vanishing remainder of the algorithm, is $r_{n-1} = \gcd(a, b)$?

Answer to Question (1). Suppose that the procedure never terminates. Then at each step we would obtain a positive remainder. The first remainder r_1 is a positive integer less than a, the second remainder r_2 is a positive integer less than r_1, etc. Therefore we have an infinite sequence of positive integers, each strictly less than the previous one:

$$a > r_1 > r_2 > r_3 \cdots > 0$$

But this is impossible. Therefore the procedure must end after a finite number of steps with a vanishing remainder.

The following lemma is needed in the answer to the second question.

4.9.1. Lemma. Let $0 < s < t$ be integers and $t = qs + r$. Then $\gcd(s, t) = \gcd(r, s)$.

Proof. Let $u = \gcd(s, t)$. Then $u|s$ and $u|(t - qs)$; hence $u|r$ and $u|s$. It remains to show that u is divisible by all common divisors of r and s. Suppose v is a common divisor of r and s. Then $v|s$ and $v|(qs + r)$; hence $v|s$ and $v|t$. Since $u = \gcd(s, t)$, we must have $v|u$. Therefore $u = \gcd(r, s)$. ■

Answer to Question (2). Let r_n be the first vanishing remainder of the procedure. Then tracing the steps backwards, we have

$$\begin{aligned}
r_{n-2} &= q_n r_{n-1} + 0 \\
r_{n-3} &= q_{n-1} r_{n-2} + r_{n-1}, \\
r_{n-4} &= q_{n-2} r_{n-3} + r_{n-2}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
r_1 &= q_3 r_2 + r_3 \\
a &= q_2 r_1 + r_2 \\
b &= q_1 a + r_1
\end{aligned}$$

Applying Lemma 4.9.1 to these equations, we obtain

$$r_{n-1} = \gcd(0, r_{n-1}) = \gcd(r_{n-1}, r_{n-2})$$
$$\gcd(r_{n-1}, r_{n-2}) = \gcd(r_{n-2}, r_{n-3})$$
$$\gcd(r_{n-2}, r_{n-3}) = \gcd(r_{n-3}, r_{n-4})$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\gcd(r_3, r_2) = \gcd(r_2, r_1)$$
$$\gcd(r_2, r_1) = \gcd(r_1, a)$$
$$\gcd(r_1, a) = \gcd(a, b)$$

Therefore $r_{n-1} = \gcd(a, b)$. ∎

As the answers to both Question (1) and Question (2) are in the affirmative, the proposed division algorithm is proved to be effective.

4.9.2. *Example.* Find gcd(12345, 56789)

Solution. The steps of the division algorithm are as follows:

$56789 = 4 \times 12345 + 7409$ $\quad (b = q_1 a + r_1)$
$12345 = 1 \times 7409 + 4936$ $\quad (a = q_2 r_1 + r_2)$
$7409 = 1 \times 4936 + 2473$ $\quad (r_1 = q_3 r_2 + r_3)$
$4936 = 1 \times 2473 + 2463$ $\quad (r_2 = q_4 r_3 + r_4)$
$2473 = 1 \times 2463 + 10$ $\quad (r_3 = q_5 r_4 + r_5)$
$2463 = 246 \times 10 + 3$ $\quad (r_4 = q_6 r_5 + r_6)$
$10 = 3 \times 3 + 1$ $\quad (r_5 = q_7 r_6 + r_7)$
$3 = 3 \times 1 + 0$ $\quad (r_6 = q_8 r_7)$

$\therefore$ gcd(12345, 56789) = 1 $\quad (\gcd(a, b) = r_7)$.

There is a more convenient way to carry out the algorithm by the following scheme.

1	12345	56789	4	q_2	a	b	q_1
	7409	49380			$q_2 r_1$	$q_1 a$	
1	4936	7409	1	q_4	$r_2 = a - q_2 r_1$	$b - q_1 a = r_1$	q_3
	2473	4936			$q_4 r_3$	$q_3 r_2$	
246	2463	2473	1	q_6	$r_4 = r_2 - q_4 r_3$	$r_1 - q_3 r_2 = r_3$	q_5
	2460	2463			$q_6 r_5$	$q_5 r_4$	
3	3	10	3	q_8	$r_6 = r_4 - q_6 r_5$	$r_3 - q_5 r_4 = r_5$	q_7
	3	9			$q_8 r_7$	$q_7 r_6$	
	0	1			$0 = r_6 - q_8 r_7$	$r_5 - q_7 r_6 = r_7$	

Finally the division algorithm also gives, as a by-product, sufficient numerical data for the evaluation of the integers x and y in

$$\gcd(a, b) = ax + by.$$

The following example will show how x and y can be expressed in terms of the successive quotients $q_1, q_2, \ldots$.

4.9.3. *Example.* Suppose the division algorithm on a pair of integers a and b terminates at step 5. Express x and y in terms of q_1, q_2, q_3, q_4 so that

$$\gcd(a, b) = ax + by.$$

Solution. Since $r_5 = 0$. We have

$$\begin{aligned}
b &= q_1 a + r_1 && \text{or} \quad r_1 = b - q_1 a \\
a &= q_2 r_1 + r_2 && \text{or} \quad r_2 = a - q_2 r_1 \\
r_1 &= q_3 r_2 + r_3 && \text{or} \quad r_3 = r_1 - q_3 r_2 \\
r_2 &= q_4 r_3 + r_4 && \text{or} \quad r_4 = r_2 - q_4 r_3 \\
r_3 &= q_5 r_4 .
\end{aligned}$$

Then $\gcd(a, b) = r_4$ and we use the above equations to eliminate the remainders r_3, r_2, r_1 successively:

$$\begin{aligned}
\gcd(a, b) = r_4 &= r_2 - q_4 r_3 \\
&= r_2 - q_4 (r_1 - q_3 r_2) \\
&= (1 + q_3 q_4) r_2 - q_4 r_1 \\
&= (1 + q_3 q_4)(a - q_2 r_1) - q_4 r_1 \\
&= a(1 + q_3 q_4) + (-q_2 - q_4 - q_2 q_3 q_4) r_1 \\
&= a(1 + q_3 q_4) + (-q_2 - q_4 - q_2 q_3 q_4)(b - q_1 a) \\
&= a(1 + q_3 q_4 + q_1 q_2 + q_1 q_4 + q_1 q_2 q_3 q_4) \\
&\quad + b(-q_2 - q_4 - q_2 q_3 q_4).
\end{aligned}$$

Thus $x = 1 + q_3 q_4 + q_1 q_2 + q_1 q_4 + q_1 q_2 q_3 q_4$

$y = -(q_2 + q_4 + q_2 q_3 q_4)$ ■

4.9.4. *Example.* Express $\gcd(289, 37)$ in the form of $289x + 37y$.

Solution. Successive divisions yield:

$$\begin{aligned}
289 &= 7 \times 37 + 30; \\
37 &= 1 \times 30 + 7; \\
30 &= 4 \times 7 + 2; \\
7 &= 3 \times 2 + 1; \\
2 &= 2 \times 1.
\end{aligned}$$

Thus for $a = 37, b = 289$, we get gcd(37, 289) = 1. The successive quotients are $q_1 = 7, q_2 = 1, q_3 = 4, q_4 = 3, q_5 = 2$. Therefore $x = 125$ and $y = -16$.
i.e. gcd(289, 37) = 289 (−16) + 37 (125) = 1. ■

4.10. EXERCISE

1. For any positive integers a and b, prove that lcm $(a, b) = ab$ if and only if a and b are relatively prime.

2. Given a positive integer k, and non-zero integers m and n, show that lcm $(km, kn) = k$ lcm (m, n).

3. Given positive integers m and n, show that gcd (m, n) = lcm (m, n) if and only if $m = n$.

4. Given positive integers m and n, show that there exist integers a and b such that gcd $(a, b) = m$ and lcm $(a, b) = n$ if and only if $m \mid n$.

*5. If the l.c.m. of two numbers is equal to the square of their difference, prove that the g.c.d. of the two numbers is the product of two consecutive integers.

6. For any positive integers a, b and m, prove that m = lcm (a, b) if and only if the following conditions are satisfied:
   (i) $a \mid m$ and $b \mid m$,
   (ii) if $a \mid c$ and $b \mid c$, then $m \mid c$.

*7. (a) If two integers m and n are relatively prime, show that mn and $m + n$ are relatively prime.
   (b) The sum of two integers is 216 and their l.c.m. is 480. Determine the two integers.

*8. (a) If m denotes the l.c.m. of the integers $a_1, a_2, \ldots, a_n$, prove that an integer N is a common multiple of $a_1, a_2, \ldots, a_n$ if and only if $m \mid N$.

(b) Given $n \in \mathbb{N}$ define $n\mathbb{Z} = \{nx : x \in \mathbb{Z}\}$.
Is it true that given $n_1, n_2 \in \mathbb{N}$ there exists $m \in \mathbb{N}$ with $n_1\mathbb{Z} \cap n_2\mathbb{Z} = m\mathbb{Z}$? Explain.
Is it true that, for some $m \in \mathbb{N}$, $n_1\mathbb{Z} \cup n_2\mathbb{Z} = m\mathbb{Z}$? Why?

9. Find gcd (272, 1479).

10. Find lcm (306, 657).

11. Let a and b be integers, not both zero. If gcd $(a, b) = d$, prove that

$$\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1.$$

12. Let a, b and c be integers. If $a \mid c$ and $b \mid c$ with gcd $(a, b) = 1$, prove that $ab \mid c$.

13. Use the division algorithm to obtain integers m and n satisfying

$$\gcd(56, 72) = 56m + 72n.$$

*14. (a) If $D = \gcd(n_1, n_2)$ where n_1 and n_2 are integers, prove that an integer N is a common divisor of n_1 and n_2 if and only if $N \mid D$.

(b) If $a_1, a_2, \ldots, a_n$ are given non-zero integers, and if

$$\begin{aligned} d_1 &= a_1, \\ d_2 &= \gcd(d_1, a_2), \\ d_3 &= \gcd(d_2, a_3), \\ &\cdots\cdots \\ d_n &= \gcd(d_{n-1}, a_n), \end{aligned}$$

and $d = \gcd(a_1, a_2, a_3, \ldots, a_n)$,

prove that $d = d_n$ and deduce that there are integers $m_1, m_2, \ldots, m_n$ such that $d = \sum_{i=1}^{n} m_i a_i$.

*15. Prove that gcd $(mn_1, mn_2) = m$ gcd (n_1, n_2) for any positive integers m, n_1 and n_2.

4.11. PRIME NUMBERS

A *prime number* (or simply a *prime*) is a positive number $p > 1$ which has no positive divisor other than 1 and p. An integer greater than 1, which is not a prime, is called a *composite number* (or simply a *composite*). It follows from the definitions that any positive number greater than 1 is either a prime or a composite. If it is a composite, then it is a product of two factors, each of which is strictly less than the given number. Should either of these factors be a composite, it can be further factorized into a product of even smaller factors. It is plausible that this process of factorization can be carried on until the given number is written as a product of primes. In other words, every positive integer greater than 1 is a product of prime numbers. It is in this sense that the prime numbers are regarded as the 'building blocks' from which other integers can be produced by multiplication. We shall see in the next section that what is briefly described above is in fact true and constitutes a part of the famous *fundamental theorem of arithmetic.* Meanwhile we shall study some general properties of prime numbers.

In our discussion of prime and composite numbers, we shall leave out the zero and the negative numbers. We take note that 1 is the only positive number that is neither a prime nor a composite. All even numbers except 2 are composite. The first ten prime numbers are 2, 3, 5, 7, 11, 13, 17, 19, 23 and 29, and the first ten composite numbers are 4, 6, 8, 9, 10, 12, 14, 15, 16 and 18.

4.11.1. Theorem. *Let p be a prime and a an integer. Then either $p = \gcd(p, a)$ or $1 = \gcd(p, a)$.*

Proof. Given p and a, it is either $p \mid a$ or $p \nmid a$. In the former case, 1 and p are the only positive common divisors of p and a. Therefore, $p = \gcd(p, a)$. In the latter case, 1 is the only common divisor of p and a. Therefore, $1 = \gcd(p, a)$. ∎

4.11.2. Theorem. *If p is a prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. For p and a, either $p \mid a$ or $p \nmid a$. If it is the former case, then there is no need to go further. So let us assume that $p \nmid a$. Then by Theorem

4.11.1, $\gcd(p, a) = 1$. Now by Example 4.6.4, it follows from $p \mid ab$ and $\gcd(p, a) = 1$ that $p \mid b$. ∎

4.11.3. Theorem. *If p is a prime and $p \mid a_1\ a_2 \ldots a_n$, then $p \mid a_i$ for at least one $i = 1, 2, \ldots n$.*

Proof. We shall prove this theorem by induction on the number n of factors. For $n = 1$, there is nothing to be proved. Suppose that the theorem holds for some $r \geqslant 1$. If $p\ a_1\ a_2 \ldots a_r\ a_{r+1}$, then it follows from Theorem 4.11.2 and $a_1\ a_2 \ldots a_r\ a_{r+1} = (a_1\ a_2 \ldots a_r)\ a_{r+1}$ that either $p \mid a_{r+1}$ or $p \mid a_1\ a_2 \ldots a_r$. If it is the former case, then there is nothing more to be proved. If it is the latter case, then $p \mid a_i$ for some $i = 1, 2, \ldots, r$ by the induction assumption. Therefore the theorem holds for any number n of factors. ∎

A special case of Theorem 4.11.3 where all the factors a_i are themselves prime numbers is worth our attention.

4.11.4 Corollary. *If p is a prime and $p \mid q_1\ q_2 \ldots q_n$ where $p, q_1, q_2, \ldots, q_n$ are all prime numbers then $p = q_i$ for at least one $i = 1, 2, \ldots, n$.*

Proof. It follows from Theorem 4.11.3 that $p \mid q_i$ for some i. Since q_i is itself a prime numbers it is divisible only by 1 and q_i. Therefore it is either $p = 1$ or $p = q_i$. But the former is impossible since p is a prime. Therefore $p = q_i$. ∎

4.11.5. Corollary. *If $p_1\ p_2 \ldots p_m = q_1\ q_2 \ldots q_n$, where the p_i and q_i are all prime numbers, then $m = n$, and after a suitable renumbering of the factors q_i, $p_i = q_i$ for all $i = 1, 2, \ldots, m$.*

Proof. We propose to prove the corollary by induction on the number m of factors. For $m = 1$, it follows from $p_1 = q_1\ q_2 \ldots q_n$ and p_1 being a prime number, that $n = 1$. Hence $p_1 = q_1$. Assume for some $r \geqslant 1$, if $p_1\ p_2 \ldots p_r = q_1\ q_2 \ldots q_n$ where the p_i and q_i are prime numbers, then $r = n$, and after a suitable renumbering, $p_i = q_i$ $(i = 1, 2, \ldots, r)$. Suppose $p_1\ p_2 \ldots p_r\ p_{r+1} = q_1\ q_2 \ldots q_n$. Then $p_{r+1} \mid q_1\ q_2 \ldots q_n$. By Corollary 4.11.4, $p_{r+1} = q_i$ for some $i = 1, 2, \ldots, n$. After a suitable renumbering, we may assume that $p_{r+1} = q_n$. Then $p_1\ p_2 \ldots p_r = q_1\ q_2 \ldots q_{n-1}$. By the

induction assumption, $r = n - 1$, and $p_i = q_i$ for $i = 1, 2, \ldots, r$ after yet another suitable renumbering of the indices of the factors $q_1, q_2, \ldots q_{n-1}$. Therefore $r + 1 = n$ and $p_i = q_i$ for $i = 1, 2, \ldots, r + 1$. Hence the corollary holds for all natural numbers m and n. ∎

4.12. THE FUNDAMENTAL THEOREM OF ARITHMETIC

The justification in regarding prime numbers as 'building blocks' of numbers is contained in the following theorem. This famous theorem can be found in Euclid's *Elements*.

4.12.1. Theorem. *The fundamental theorem of arithmetic. Every integer greater than 1 can be written as a product of prime numbers and such representation is unique up to the order of the prime factors.*

Proof. The theorem consists of two parts:

(a) Existence. Given any integer $a > 1$, there are $n \geqslant 1$ prime numbers $p_1, p_2, \ldots, p_n$, not necessarily all distinct, such that

$$a = p_1 p_2 \ldots p_n$$

(b) Uniqueness. If $a = p_1 p_2 \ldots p_n$ and $a = q_1 q_2 \ldots q_m$ where p_i and q_i are all prime numbers, then $n = m$, and after a suitable renumbering of the indices, $p_i = q_i$ for $i = 1, 2, \ldots, n$.

The second part of the theorem has been established in Corollary 4.11.5. Therefore we need only show the existence of such prime numbers p_i.

Suppose to the contrary, that there are positive integers greater than 1 that are not expressible as a product of prime numbers. Then by the well-ordering principle, there is a least integer among them. Let this be c. Then by definition, c must have the following properties:

(i) $c > 1$;

(ii) c is not a product of prime numbers;

(iii) If d is an integer such that $1 < d < c$, then d is a product of prime numbers.

If follows from (ii) that c itself cannot be a prime number. Therefore we can find two positive integers d_1 and d_2, such that $c = d_1 d_2$ and $1 < d_i < c$ $(i = 1, 2)$. By (iii) both d_1 and d_2 can be written as a product of prime numbers. Therefore c itself is a product of prime numbers contradicting

(ii). Hence the assumption that there exist integers failing to be a product of prime numbers must be wrong. The proof is now complete. ■

4.12.2. Addendum. The uniqueness part of the theorem can be put in a slightly improved formulation, if we arrange the prime factors in the order of magnitude. Denote by $p_1, p_2, p_3, \ldots$ the sequence of all prime numbers 2, 3, 5, 7, . . . in the order of increasing magnitude. The fundamental theorem of arithmetic can be formulated in the following way:

For every integer a greater than 1, there exists a unique set of natural numbers $e_1, e_2, \ldots, e_n$ *such that*

$$a = p_1^{e_1} p_2^{e_2} \ldots p_n^{e_n}$$

where $e_n > 0$.

For example, 7500038 = 2 x 7 x 7 x 7 x 13 x 29 x 29. Therefore $7500038 = 2^1 \times 3^0 \times 5^0 \times 7^3 \times 11^0 \times 13^1 \times 17^0 \times 19^0 \times 23^0 \times 29^2$.

4.12.3. *Remarks.* The fundamental theorem tells us that any number can be factorized in essentially one way as a product of prime numbers, but it contains no information as to how a given number can be factorized. In fact the factorization of large numbers remains one of the most difficult problems of mathematics. In recent years, great improvement has been made in the construction of very high speed computers and much progress has been achieved in the field of algorithms. It is now possible to factorize some very large numbers and this has opened up many applications of the theory of numbers. Even so, using the best algorithms, it still may take 10 minutes to determine the primality of a 200-digit integer on the best computer available. The factorization of a 100-digit composite number may even take months to complete. But by the primitive method of trial and error, it may take many years to find the prime factors of a 50-digit number on a computer which could perform one division in one billionth of a second.

4.13. THE INFINITY OF PRIME NUMBERS

Besides the fundamental theorem of arithmetic, Euclid has also proved the following classical theorem on the supply of prime numbers.

4.13.1. Theorem. *There is an infinite supply of prime numbers.*

Proof. Suppose to the contrary that the theorem is false, and there were only a finite number of prime numbers. Denote these primes by $p_1, p_2, \ldots, p_n$ and consider the positive integer

$$c = p_1 p_2 \ldots p_n + 1$$

Since $p_1, p_2, \ldots, p_n$ are all the prime numbers and $c \neq p_j$ for all $j = 1, 2, \ldots, n$, the number c must be a composite number. By the fundamental theorem of arithmetic, c must be divisible by some prime number, say p_j for some $j = 1, 2, \ldots, n$. Then it follows from $p_j \mid c$ and $p_j \mid p_1 p_2 \ldots p_n$ that $1 = c - p_1 p_2 \ldots p_n$ is also divisible by the prime number p_j. But this is clearly absurd. Therefore it cannot be the case that there are only a finite number of primes. ∎

For over two thousand years, mathematicians all over the world have been working to extend the list of prime numbers 2, 3, 5, 7, 11, ... Eratosthenes of Cyrene (276 – 194 B.C.) was probably the first to have found a systematic method. The method is known as the *Sieve of Eratosthenes* and is based on the following property of composite numbers.

4.13.2. Lemma. *Let a be an integer greater than 1. If a is composite then it is divisible by a prime number p not exceeding $\sqrt{a}$.*

Proof. Let $a = bc$ where $1 < b < a$ and $1 < c < a$. Suppose that $b \leqslant c$, then $b^2 \leqslant bc = a$, and so $b \leqslant \sqrt{a}$. Since $b > 1$, b must be divisible by a prime number p by Theorem 4.12.1. It follows from $p \mid b$ that $p \leqslant b \leqslant \sqrt{a}$, and $p \mid a$ since $a = bc$. ∎

It follows from the above lemma that if $a > 1$ is not divisible by any prime number $p \leqslant \sqrt{a}$, then a is itself a prime number. Based on this we have

4.13.3. Sieve of Eratosthenes for finding all prime numbers between 1 and a given number n. The method consists of the following steps. To start, we write down all integers from 1 to n in a block formation.

1	2	...	...	m
$m+1$	$m+2$	...	...	$2m$
$2m+1$	$2m+2$	...	...	$3m$
...	...	...	...	...
...	...	$n-1$	n	

where m is the least integer which is not less than $\sqrt{n}$. Then we proceed to pick up all prime numbers and eliminate all composite numbers of the block by a systematic search to be carried out in several rounds.

Round 1. Cross out the number 1.
Round 2. The next number is 2 which is a prime number. Put a circle around the number 2 and cross out all higher multiples (i.e. every second number) 4, 6, . . . , $2t$, . . . of the prime number 2 in the block.
Round 3. The next untreated (i.e. neither circled nor crossed out) number is 3. Because 3 is not divisible by any prime number less than itself, it must be a prime number. Circle 3 and cross out all higher multiples (i.e. every third number) 3, 6, 9, . . . $3t$, . . . of 3 in the block.
Round 4. The next untreated number is 5. The number 5, not being crossed out in the previous rounds, is not divisible by any prime number less than itself; it must be a prime number. Circle 5 and cross out all higher multiples (i.e. every fifth number) 10, 15, . . . , $5t$, . . . of 5 in the block.
Round 5. . . . . . .
. . . . . .

The process terminates as soon as all numbers on the first row of the block are either circled or crossed out. Then all numbers in the block that are crossed out are composite numbers. The remaining numbers are then all the prime numbers from 1 to n. These are numbers that fall through the 'sieve'. In particular the circled numbers are the primes $p \leqslant \sqrt{n}$ while the others are numbers not divisible by any prime $p \leqslant \sqrt{n}$.

The table below is the result of a sieving process. The multiples of 2 are crossed out by $-$; those of 3 by $|$; those of 5 by \; and those of 7 by /.

~~1~~	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	~~7~~	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~

Thus the prime numbers less than 101 are:
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

Obviously the sieve of Eratosthenes is not a suitable method for finding very large prime numbers. Among the many mathematicians who have contributed to the study of prime numbers, a French monk, Marin Mersenne (1588 – 1648) suggested that prime numbers can be found among numbers of the form

$$M_p = 2^p - 1$$

where p is a prime number. Prime numbers of this kind are now called *Mersenne primes*. He claimed that for $p = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127$ and 257, M_p is a prime and that M_p is composite for all other primes $p < 257$. Given the limitation of calculating technique of his day, it is impossible that he has actually verified his claim except possibly the first few on his list: $M_2 = 3$, $M_3 = 7$, $M_5 = 31$, $M_7 = 127$, $M_{13} = 8191$, $M_{17} = 131071$ and $M_{19} = 524287$. Indeed it was more than a hundred years later that the Swiss mathematician Leonard Euler (1707 – 1783) verified that $M_{31} = 2147483647$ is a prime number. A complete answer to Mersenne's conjecture was finally available in 1947: M_{61}, M_{89} and M_{107} are also prime numbers while M_{67} and M_{257} turned out to be composite numbers. Up to now 30 Mersenne primes have been found. The last one M_{216091} was discovered in September 1985 and is a number of 65050 digits.

4.14. EXERCISE

1. Prove that one can always find n consecutive composite integers, however great n may be.

2. If $p_1, p_2, \ldots, p_n$ are n distinct prime numbers, prove that the sum $\frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_n}$ is not an integer.

3. If m and n are integers greater than 1, prove that $m^4 + 4n^2$ is a composite number.

4. A perfect number is a positive integer n which is the sum of all its positive factors other than itself, e.g. $6 = 1 + 2 + 3$ is the smallest perfect number. If $p = 2^{n+1} - 1$ is prime, prove that $2^n p$ is a perfect number.

5. Let p be the smallest prime factor of an integer n, and suppose that $p > \sqrt[3]{n}$. Show that $\frac{n}{p}$ is also prime.

6. Prove that every odd prime can be expressed as the difference of two squares in one and only one way.

7. 3, 5 and 7 are prime numbers. Are there any other triplets of prime numbers of the form n, $n + 2$ and $n + 4$? Why?

*8. Prove that $\binom{n}{r}$ is never prime for any positive integers r and n such that $2 \leqslant r \leqslant n - 2$.

*9. If p is a prime number not less than 5, prove that $p^2 - 1$ is divisible by 24.

10. (a) If n is a composite number, show that $2^n - 1$ is composite.
    (b) If $2^n - 1$ is a prime number, show that n must also be prime.

11. (a) If a, b and c are integers and the equation $ax^2 + bx + c = 0$ has a rational root m/n where m and n are relatively prime integers, show that m and n are factors of c and a respectively.
    (b) Hence show that $\sqrt{p}$ is irrational if p is a prime number.

12. A way of determining whether a given number is prime or not is shown below:

    Let N be a given number, and suppose that n is the least number for which $n^2 \geqslant N$. Form the numbers

    $$n^2 - N, (n + 1)^2 - N, (n + 2)^2 - N, \ldots$$

    until one of these is a perfect square. Let the perfect square reached be m^2. Then

    $$(n + i)^2 - N = m^2$$

    $$\Leftrightarrow$$

    $$N = (n + i + m)(n + i - m).$$

(a) Determine whether 8453 is a prime number.

*(b) What happens when N is a prime number?

*13. If $2^n + 1$ is a prime number, prove that n is a power of 2.

*14. If p_n is the nth prime number, prove that $p_n \leqslant 2^{2^{n-1}}$.

*15. If p is a prime number and a, d are positive integers not divisible by p, prove that exactly one of the numbers

$$a + d,\ a + 2d,\ a + 3d,\ \ldots,\ a + (p-1)d$$

is divisible by p.

16. If p is a prime number, prove that

(a) $\binom{p}{r}$ is divisible by p for $r = 1, 2, 3, \ldots, p-1$, and

*(b) $\binom{p-1}{r} - (-1)^r$ is divisible by p for $r = 0, 1, 2, \ldots, p-1$.

17. (a) m and n are positive integers such that $m + n$ and $m^2 + n^2$ are both divisible by an odd prime p. Prove that m and n are both divisible by p.

(b) If p is a prime number greater than 3 and a, b, c are positive integers such that p divides each of

$$a + b + c,\ a^2 + b^2 + c^2 \text{ and } a^3 + b^3 + c^3.$$

Prove that p divides each of a, b and c.

*18. If m and n are integers with no common factors other than 1, prove that there are integers a and b such that

$$\frac{1}{mn} = \frac{a}{m} + \frac{b}{n}.$$

Deduce from this that every rational number can be expressed in the form

$$\frac{a_1}{q_1} + \frac{a_2}{q_2} + \ldots + \frac{a_n}{q_n}$$

where the numerators $a_1, a_2, \ldots, a_n$ and the denominators $q_1, q_2, \ldots, q_n$ are integers and each denominator is a prime number or a power of a prime number.

19. If $a^k - 1$, where k is a positive integer, is a prime number, prove that $a = 2$.

*20. Prove that there are infinitely many prime numbers of the form $3n + 2$, n being a natural number.

21. (a) For any positive integer N, prove that the l.c.m. of the numbers 1, 2, 3, . . . , N can be expressed as $2^n M$, where M is an odd number and n is the greatest integer such that $2^n \leqslant N$.
    (b) Hence prove that, if $N > 1$, then the sum
    $$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{N}$$
    can be expressed as $\frac{a}{b}$ where a is odd and b is even.

22. Suppose that a natural number N can be factorized with the primes p_1, p_2, . . . , p_r, each raised to the powers a_1, a_2, . . . , a_r respectively, i.e. $N = {p_1}^{a_1} {p_2}^{a_2} \ldots {p_r}^{a_r}$. Find an expression for the number of positive divisors of N.

4.15. CONGRUENCE

While the topics of the previous sections of this chapter are discussed in Euclid's *Elements*, the subject matter of the rest of the chapter is developed by the German mathematician Carl Friedrich Gauss (1777 – 1855) in his book *Disquisitiones Arithmeticae*, which is often regarded as the first book on modern number theory. Let us first put down the definition of congruence as given by Gauss.

4.15.1. Definition. *Let n be a fixed integer greater than 1. Two integers a and b are said to be congruent modulo n if $a - b$ is divisible by n.* In this case we write $a \equiv b \pmod{n}$; otherwise we write $a \not\equiv b \pmod{n}$.

For example, $11 \equiv 32 \pmod{7}$; $-7 \equiv 25 \pmod{8}$; $2 \not\equiv 19 \pmod{3}$; $7 \not\equiv -3 \pmod{6}$. The trivial case of congruence modulo 1 is intentionally excluded.

One obvious relationship between congruence and division is that if r

is the remainder of a when divided by n (i.e. $a = qn + r$), then $a \equiv r \pmod{n}$. In other words, in the theory of congruence modulo n, no distinction shall be made between a and its remainder r when divided by n. Thus we may view the theory of congruence as the arithmetic of the remainders. For example, the integers $1, 7, 13, \ldots$ and $-5, -11, -17, \ldots$ will become indistinguishable reduced modulo 6. In particular both 5721745 and -15405713 can be replaced by 1 and

$$5721745 \times (-15405713) \equiv 1 \times 1 \equiv 1 \pmod{6}.$$

The basic rules of congruence are listed in the theorem below.

4.15.2. Theorem. *Let n be a fixed positive integer and let a, b, c, d be four arbitary integers. Then the following statements hold:*

(a) $a \equiv a \pmod{n}$;

(b) *If* $a \equiv b \pmod{n}$, *then* $b \equiv a \pmod{n}$;

(c) *If* $a \equiv b \pmod{n}$ *and* $b \equiv c \pmod{n}$, *then* $a \equiv c \pmod{n}$;

(d) *If* $a \equiv b \pmod{n}$ *and* $c \equiv d \pmod{n}$, *then* $a + c \equiv b + d \pmod{n}$ *and* $ac \equiv bd \pmod{n}$;

(e) *If* $ac \equiv bc \pmod{n}$ *and if c and n are relatively prime, then* $a \equiv b \pmod{n}$.

Rules (a), (b) and (c) govern the general usage of the congruence symbol $\equiv$. From rule (d) the usual commutative, associative and distributive laws follow. Rule (e) is the cancellation law where the condition $\gcd(c, n) = 1$ takes the place of $c \neq 0$ in the ordinary cancellation law: if $ac = bc$ and if $c \neq 0$ then $a = b$. The proof of these rules is left as an exercise for the reader.

Let us carry the similarity between congruence and equality another step further and consider the counterpart of a first-degree equation

$$ax = b$$

in the unknown x with integral coefficients a and b. Such equation has no integral solution unless b is divisible by a. Substituting $=$ by $\equiv \pmod{n}$ we obtain a *linear congruence*

$$ax \equiv b \pmod{n}$$

For the solution of linear congruence we have the following theorem.

4.15.3. Theorem. *If* $\gcd(a, n) = 1$, *then the linear congruence* $ax \equiv b \pmod{n}$ *has integral solutions. Moreover, any two solutions are congruent modulo n.*

Proof. It follows from $\gcd(a, n) = 1$ that we can find integers x_0 and y_0 such that $1 = ax_0 + ny_0$. Therefore $b = a(bx_0) + n(by_0)$; and hence $a(bx_0) \equiv b \pmod{n}$. Thus $x = bx_0$ is a solution of the given linear congruence. Suppose that x and x' are both solutions. Then it follows from $ax \equiv b \pmod{n}$ and $ax' \equiv b \pmod{n}$ that $ax \equiv ax' \pmod{n}$. Because a and n are relatively prime, we may apply the cancellation law to get $x \equiv x' \pmod{n}$.■

4.16. CHINESE REMAINDER THEOREM

In the ancient Chinese mathematics text 孫子算經 *Sun Zi Suan Jing*, one finds the following problem on the solution of a system of linear congruences.

今有物不知其數，三三數之賸二，五五數之
賸三，七七數之賸二，問物幾何？

(Find a number which leaves the remainders 2, 3, 2 when divided by 3, 5, 7 respectively),

i.e. find x such that
$$x \equiv 2 \pmod{3}$$
$$x \equiv 3 \pmod{5}$$
$$x \equiv 2 \pmod{7}$$

A solution to the problem is given in the text as

$$2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 233$$

which gives 23 when reduced by multiples of 105.

Similar problems are also found in ancient Greek mathematics. The general solution of a system of linear congruences is now known as the Chinese remainder theorem.

4.16.1. Chinese remainder theorem. *Let $n_1, n_2, \ldots, n_r$ be positive integers such that they are pairwise relatively prime, i.e.* $\gcd(n_i, n_j) = 1$ *for* $i \neq j$. *Then the system of linear congruences*

$$x \equiv b_1 \pmod{n_1}$$
$$x \equiv b_2 \pmod{n_2}$$
$$\ldots\ldots\ldots\ldots$$
$$x \equiv b_r \pmod{n_r}$$

has solutions. Furthermore any two solutions are congruent modulo $n_1 n_2 \ldots n_r$.

Proof. Let $n = n_1\, n_2 \ldots n_r$ and let $m_k = n/n_k = n_1 \ldots n_{k-1}\, n_{k+1} \ldots n_r$ for $k = 1, 2, \ldots, r$. Since $\gcd(n_i, n_j) = 1$, n_i and n_j $(i \neq j)$ have no common prime factors. Therefore n_k and m_k have no common prime factors, so $\gcd(m_k, n_k) = 1$ for $k = 1, 2, \ldots, r$. By Theorem 4.15.3, each congruence $m_k x \equiv 1 \pmod{n_k}$ has solutions. Let x_k be any one of these solutions. Then

$$m_k\, x_k = 1 \pmod{n_k} \text{ for } k = 1, 2, \ldots, r.$$

Consider the integer $b = b_1\, m_1\, x_1 + b_2\, m_2\, x_2 + \ldots + b_r\, m_r\, x_r$.
Then it follows from $m_i \equiv 0 \pmod{n_k}$ for $i \neq k$, that

$$b \equiv b_k\, m_k\, x_k \equiv b_k\, 1 \equiv b_k \pmod{n_k} \text{ for } k = 1, 2, \ldots, r.$$

Thus b is a solution of the system.

Suppose that b and b' are two such solutions. Then it follows from $b \equiv b_k \equiv b' \pmod{n_k}$ that $n_k \mid (b - b')$ for $k = 1, 2, \ldots, r$. Hence $n_1\, n_2 \ldots n_k \mid (b - b')$ since n_i and n_k have no common prime factors for $i \neq k$. Therefore $b \equiv b' \pmod{n_1\, n_2 \ldots n_k}$. The proof is now complete. ∎

4.17. EXERCISE

1. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, prove that
   (i) $a + c \equiv b + d \pmod{n}$,
   (ii) $a - c \equiv b - d \pmod{n}$,
   (iii) $ac \equiv bd \pmod{n}$.

2. For arbitrary integers a and b, prove that $a \equiv b \pmod{n}$ if and only if a and b leave the same non-negative remainder when divided by n.

3. Solve for x : $75x \equiv 2 \pmod{13}$.

4. Solve for x : $43x \equiv 5 \pmod{61}$.

5. Prove that the indeterminate equation
   $$2x - 6y = 2n + 1$$
   has no integral solutions when n is any natural number.

6. Prove that a positive integer is divisible by 9 if and only if the sum of its digits is divisible by 9.

7. If $ma \equiv mb \pmod{n}$ where m is an integer, show that
$$a \equiv b \pmod{\frac{n}{\gcd(m, n)}}.$$

8. Show that the linear congruence $ax \equiv b \pmod{m}$ has a root if and only if $\gcd(a, m) \mid b$.

9. Solve for x and y the congruences
$$\begin{cases} 4x - 5y \equiv 1 \pmod{7}, \\ x + 2y \equiv 3 \pmod{7}. \end{cases}$$

10. (a) If $11n \equiv 37 \pmod{111}$, show that n is a multiple of 37.
    (b) If the number $100a + 10b + c$ is divisible by 37, prove that $100b + 10c + a$ and $100c + 10a + b$ are also divisible by 37.

*11. (a) If p is a prime number, show that all the binomial coefficients, except the first and the last, in the expansion of $(1 + x)^p$, are divisible by p.
    (b) Prove that, for every positive integer n and every prime number p, $n^p \equiv n \pmod{p}$. Hence deduce that $n^{p-1} \equiv 1 \pmod{p}$ for all integers n not divisible by p (Fermat's theorem).

*12. Let n be a fixed positive integer and a, b be arbitrary integers.
    (a) If $a \equiv b \pmod{n}$, prove that $a^k \equiv b^k \pmod{n}$ for any positive integer k.
    (b) Let $p(x) = \sum_{k=0}^{m} C_k x^k$ be a polynomial in x with integral coefficients C_k $(k = 0, 1, 2, \ldots, m)$.
    If $a \equiv b \pmod{n}$, prove that $p(a) \equiv p(b) \pmod{n}$.
    (c) Let $N = \sum_{k=0}^{m} a_k \cdot 10^k$ $(0 \leqslant a_k < 10)$ be the decimal representation of the positive integer N and let $T = \sum_{k=0}^{m} (-1)^k a_k$. Prove that $11 \mid N$ if and only if $11 \mid T$.

5. *The Real Numbers*

5.1. THE NUMBER LINE $\mathbb{R}$

For the purpose and the scope of this course, it will be sufficient for us to think of the real numbers as the points along a straight line which extends indefinitely in both directions. (Fig. 5.1) This straight line is referred to as the *number line*. Any finite portion of the number line can

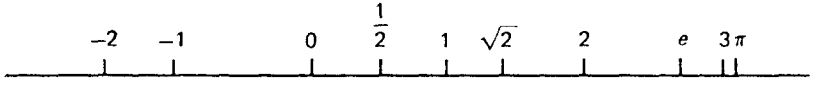


Fig. 5.1

serve as a ruler with which we may measure the lengths of line segments in Euclidean geometry. We shall denote the set of all real numbers by the bold faced capital $\mathbb{R}$. The three most important subsets of $\mathbb{R}$ are listed in the table below:

Subset	*Notation*	*Elements*
Natural numbers	$\mathbb{N}$	$0, 1, 2, \ldots$
Integers	$\mathbb{Z}$	$\ldots -2, -1, 0, 1, 2, \ldots$ i.e. $\pm n$ where $n \in \mathbb{N}$.
Rational numbers	$\mathbb{Q}$	$0, 1, 2, -1, \frac{1}{2}, -\frac{3}{2}, \ldots$ i.e. a/b where $a, b \in \mathbb{Z}$ and $b \neq 0$.

Natural numbers and integers have been discussed in some detail in Chapters 2 and 4. Rational numbers are also known as *fractions*. They can be either written as

$$\frac{m}{n} \text{ or } -\frac{m}{n}, n \neq 0$$

where m, n are natural numbers that have no common factors except 1. A rational number which is expressed in this form is said to be in its *lowest*

terms. For example, $-\frac{1}{2}, -\frac{5}{3}, \frac{12345}{56789}$ are in their lowest terms whereas 12/15, 8/100 are not. Furthermore all decimal numbers, such as 2.13, −7.186, are rational numbers.

For the three subsets above, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ holds. Moreover $\mathbb{N}$ is a proper subset of $\mathbb{Z}$, for example, $-3 \notin \mathbb{N}$. It is also true that $\mathbb{Z}$ is a proper subset of $\mathbb{Q}$, for example, $3/2 \notin \mathbb{Z}$. It is less obvious that $\mathbb{Q}$ is a proper subset of $\mathbb{R}$. In fact, until the discovery of the irrationality of $\sqrt{2}$, the Pythagoreans of ancient Greece believed that any two line segments are commensurable (i.e. there are no real numbers other than the rational numbers).

5.1.1. *Example.* $\sqrt{2}$ is a real number which is not a rational number.

Proof. $\sqrt{2}$ is the length of the diagonal of a square with unit side length. The number can be located on the number line between the rational numbers 1.41 and 1.42 by the following construction (Fig. 5.2):

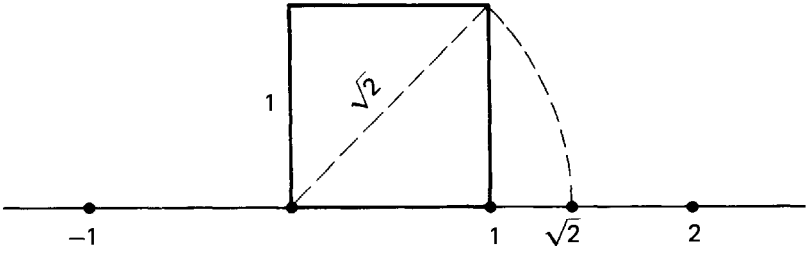


Fig. 5.2

Thus $\sqrt{2}$ is a real number.

Suppose that $\sqrt{2}$ were a rational number. Then it would be possible to find positive integers m and n such that

$$\sqrt{2} = \frac{m}{n} \quad \text{and} \quad \gcd(m, n) = 1.$$

Squaring the first equation, we obtain $2n^2 = m^2$. Therefore $2 | m^2$. By Theorem 4.11.2, we get (i) $2 | m$. Consequently $2^2 | m^2$. Since $m^2 = 2n^2$, it follows that $2^2 | 2n^2$. Hence $2 | n^2$. Thus we also have (ii) $2 | n$. But (i) and (ii) are contradictory to $\gcd(m, n) = 1$. Therefore it is not possible that $\sqrt{2}$ is a rational number. ∎

Therefore $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$ holds. Real numbers which are not rational numbers are called *irrational numbers*. $\sqrt{2}$, e, π and a great deal more are irrational numbers.

5.2. EXERCISE

1. Prove that $\sqrt[3]{2}$ is irrational.

2. For any integer $n > 1$ and any prime number p, prove that $p^{1/n}$ is irrational.

3. If r and s are rational numbers, prove that $r + s\sqrt{2}$ is irrational unless $s = 0$.

4. Read the following passage carefully.
   'Let $S = 1 - 1 + 1 - 1 + 1 - \ldots$
   Then $S = (1 - 1) + (1 - 1) + (1 - 1) + \ldots$
   $= 0 + 0 + 0 + \ldots$
   $= 0.$
   But $S = 1 - (1 - 1) - (1 - 1) - (1 - 1) \ldots$
   $= 1 - 0 - 0 - 0$
   $= 1.$
   Also $S = 1 - (1 - 1 + 1 - 1 + 1 - \ldots)$
   $= 1 - S.$
   Hence $2S = 1$, and $S = 1/2$.'
   Now which is the correct value of S? Why?

5. If a, b, c and d are rational numbers such that $a + b\sqrt{3} = c + d\sqrt{3}$, prove that $a = c$ and $b = d$.

6. If a and b are positive rational numbers and $\sqrt{a}$ is irrational, prove that $\sqrt{a} + \sqrt{b}$ is irrational.

7. Show that $\log_{10} 2$ is irrational.

8. Prove that the root of the equation $2^x = 3$ is irrational.

9. (a) If a is rational and b is irrational, is $a + b$ necessarily irrational? What if a and b are both irrational?
   (b) If a is rational and b is irrational, is ab necessarily irrational?
   (c) Is there a number a such that a^2 is irrational, but a^4 is rational?
   (d) Are there two irrational numbers whose sum and product are both rational?

10. (a) If $x = p + \sqrt{q}$ where p and q are rational numbers, and n is a positive integer, prove that $x^n = a + b\sqrt{q}$ for some rational numbers a and b.
    (b) Prove also that $(p - \sqrt{q})^n = a - b\sqrt{q}$.

11. If m and n are unequal natural numbers, prove that $\log_{10}(2^m\, 5^n)$ is irrational.

12. If $x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1 x + a_0 = 0$ for some integers a_{n-1}, $a_{n-2}, \ldots, a_1, a_0$, prove that x is irrational unless x is an integer.

*13. If n is a natural number greater than 1, prove that $\sqrt{n^2 - 1}$ is irrational.

14. Show that there is no rational number whose square is equal to 3.

15. Explain the following paradox:
'Let a/b and c/d be any two fractions representing the same rational number. Then

$$\frac{a}{b} - 1 = \frac{c}{d} - 1$$

and

$$\frac{a}{b} + 1 = \frac{c}{d} + 1.$$

It follows that $\frac{b}{d} = \frac{a-b}{c-d} = \frac{a+b}{c+d}$.

$$\therefore \quad (a-b)(c+d) = (a+b)(c-d).$$

$$\therefore \quad ad - bc = -ad + bc.$$

Hence $1 = -1$.'

5.3. SOME BASIC ASSUMPTIONS

It is not our intention to go back to the very first principle to develop formally the system $\mathbb{R}$ from a set of axioms, as this would be inappropriate

for the present course. Instead we shall adopt a straight forward approach, by accepting without question some very well-known properties of real numbers as our basic assumptions. Firstly these include all the usual laws of addition, multiplication, subtraction and division. These are then our assumptions on arithmetic. In particular the division by zero is not allowed. Thus it is meaningless, for example, to write

$$\frac{3}{0}.$$

Later on we shall specify the precise meaning of the so called 'infinity' symbol ∞ but under no circumstance are we allowed to write

$$\frac{4}{0}=\infty \quad \text{or} \quad -\frac{5}{0}=-\infty.$$

The next group of assumptions concern the natural order of real numbers. These are the following six rules of proper usage of the inequality signs $<$, $\leqslant$, $>$, and $\geqslant$.

5.3.1. Rule. *Given any two real numbers a and b there are three mutually exclusive possibilities:*

$a > b$ *(a is greater than b)*
$a = b$ *(a equals b)*
or $a < b$ *(a is less than b).*

For $a > b$ we may also write $b < a$. By $a \leqslant b$ (a is less than or equal to b, or a is not greater than b), we mean either $a < b$ or $a = b$. A stroke across an inequality sign means that the inequality does not hold. Thus, for example, $1 < 3$, $-2 \leqslant 4$, $3 \leqslant 3$, $3 \not> 5$, $-5 \not\leqslant -7$ are all true statements. So are the following statements:

$a \neq b$ if and only if $a > b$ or $b < a$
$a \not< b$ if and only if $a \geqslant b$
$a \not\leqslant b$ if and only if $b < a$.

The next rule expresses the transitivity of the order relation.

5.3.2. Rule. *If $a < b$ and $b < c$, then $a < c$.*

Then we have the next three rules concerning inequalities in conjunction with arithmetic.

5.3.3. Rule. *If $a < b$, then $a + c < b + c$ for any real number c.*

5.3.4. Rule. *If $a < b$, and $c > 0$, then $ac < bc$.*

5.3.5. Rule. *If $a < b$ and $c < 0$, then $ac > bc$.*

Finally we assume one more rule in relation to the natural numbers.

5.3.6. Archimedian postulate. *Given any two positive real numbers a and b, we can find a natural number n such that $na > b$.*

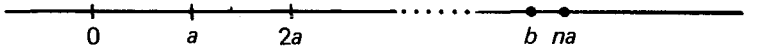


Fig. 5.3

The last rule corresponds geometrically to the familiar process of estimating the distance between two points (in this case 0 and b) by a measuring stick of length a; it guarantees that if we start at one point and lay off a succession of equal distances (the length of the measuring stick) towards the second point, we will eventually pass the second point (Fig. 5.3).

From these six rules, all the other (correct) inequalities may be derived. Let us look at a few examples.

5.3.7. *Example.* It follows from Rule 5.3.5. that if $a < b$, then $-a > -b$.

5.3.8. *Example.* Show that if $a > 0$ then $a^{-1} > 0$.

Proof. Suppose that $a > 0$ but $a^{-1} \not> 0$. Then either $a^{-1} = 0$ or $a^{-1} < 0$. If it were the case that $a^{-1} = 0$, then $1 = a^{-1}a = 0 \cdot a = 0$, which is impossible. If it were the case that $a^{-1} < 0$, then by Rule 5.3.4, $a\,a^{-1} < 0$. Hence $1 < 0$ which is also impossible. Therefore we must conclude that $a^{-1} > 0$. ■

5.3.9. *Example.* Prove that if $a < b$ and $c < d$, where b and c are both positive, then $ac < bd$.

Proof. By Rule 5.3.4, it follows from $a < b$ and $c > 0$ that $ac < bc$. Similarly, we have $bc < bd$. Therefore by Rule 5.3.2, we get $ac < bd$. ■

5.3.10. *Example.* Show that given any positive real number c and any positive integer n, we can find a decimal number a such that $|c - a| < 10^{-n}$. In other words, every real number can be approximated to any degree of accuracy by a decimal number.

Proof. If we denote the approximating decimal number a by

$$a_0.\, a_1 a_2 \ldots a_i \ldots a_n$$

where a_0 is a natural number and the decimal digits $a_1, a_2, \ldots, a_n$ have values 0, 1, 2, . . . 8 or 9, then we require this number to satisfy the condition that

$$a_0.\, a_1 a_2 \ldots a_i \leqslant c < a_0.\, a_1 a_2 \ldots a_i + 10^{-i}$$

for every $i = 0, 1, 2, \ldots, n$. Thus

$$|c - a_0.\, a_1 a_2 \ldots a_i| < 10^{-i}.$$

Now we proceed to find the integer a_0 and the digits $a_1, a_2, \ldots a_i$ one by one.

Geometrically a_0 is the integer such that the number c lies between a_0 and $a_0 + 1$. To find a_0, we apply the Archimedian postulate to the numbers 1 and c to find a positive integer m such that $c < m$. By the well-ordering principle, we should have a least positive integer $a_0 + 1$ such that $c < a_0 + 1$. Hence $0 \leqslant a_0 \leqslant c < a_0 + 1$ (Fig. 5.4).

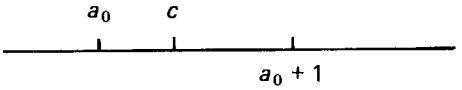


Fig. 5.4

After dividing the interval between a_0 and $a_0 + 1$ into 10 subintervals of length 10^{-1}, we shall find that c falls in one such subinterval, say between $a_0.\, a_1$ and $a_0.\, a_1 + 10^{-1}$ (Fig. 5.5).

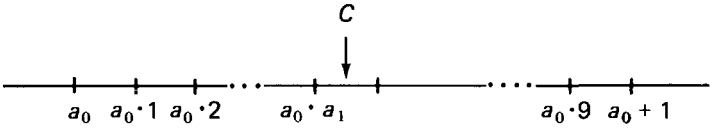


Fig. 5.5

Alternately we may apply the Archimedean postulate and the well-ordering principle to the number 10^{-1} and $c - a_0$ to get the least positive

integer $a_1 + 1$ such that $c - a_0 < (a_1 + 1)10^{-1}$. Then

$$10^{-1}a_1 \leqslant c - a_0 < (a_1 + 1)10^{-1},$$

i.e.

$$a_0.a_1 \leqslant c < a_0.a_1 + 10^{-1}$$

For the same reason we should be able to find the least integer $a_2 + 1$ such that $c - a_0.a_1 < (a_2 + 1)10^{-2}$. Then

$$a_0.a_1a_2 \leqslant c < a_0.a_1a_2 + 10^{-2}$$

Geometrically c falls in the subinterval of length 10^{-2} between $a_0.a_1a_2$ and $a_0.a_1a_2 + 10^{-2}$.

Carrying out the process another $n - 2$ steps, we get $a_0, a_1, a_2, \ldots, a_n$ such that

$$a_0.a_1a_2 \ldots a_n \leqslant c < a_0.a_1a_2 \ldots a_n + 10^{-n}$$ ■

5.4. EXERCISE

1. Given two positive real numbers a and b, prove that $a < b$ if and only if $a^2 < b^2$.

2. What has gone wrong in the following?

$$4 > 2$$
$$\therefore\ 4 \log \frac{1}{2} > 2 \log \frac{1}{2}$$
$$\therefore\ \log (\frac{1}{2})^4 > \log (\frac{1}{2})^2$$
$$\therefore\ (\frac{1}{2})^4 > (\frac{1}{2})^2$$
$$\therefore\ \frac{1}{16} > \frac{1}{4}.$$

3. Let $a < b$ and $c < d$. Determine which of the following statements hold. Why?
   (i) $a - c < b - d$;
   (ii) $a + c < b + d$;
   (iii) $\frac{a}{c} < \frac{b}{d}$;
   (iv) $ac < bd$.

4. Approximate $\sqrt{2}, \sqrt{5}$ and $\frac{11}{7}$ to the sixth decimal place.

5.5. SOME WELL-KNOWN INEQUALITIES

5.5.1. Theorem. *The inequality* $2^n > n$ *holds for all natural numbers* $n = 0, 1, 2, \ldots$.

Proof. For $n = 0$, the inequality is $2^0 = 1 > 0$ which is obviously true. Suppose that $2^k > k$ holds for some $k \geqslant 0$. Then

$$2^{k+1} = 2^k + 2^k \geqslant 2^k + 1 > k + 1.$$

Therefore by induction, $2^n > n$ holds for all n. ∎

5.5.2. Theorem. *The inequality* $2^n > n^2$ *holds for all integers* $n \geqslant 5$.

Proof. We observe first that $(k-1)^2 > 2$ for all integers $k \geqslant 5$. Hence $k^2 > 2k + 1$.

Now we proceed to prove the theorem by induction on n. For $n = 5$, $2^5 = 32 > 25 = 5^2$; the inequality is obviously true. Suppose that $2^k > k^2$ holds for some $k \geqslant 5$. Then

$$2^{k+1} = 2 \cdot 2^k > k^2 + k^2 > k^2 + 2k + 1 = (k+1)^2.$$

Therefore the theorem holds for all $n \geqslant 5$. ∎

5.5.3. Theorem. *If* $\frac{a}{b} < \frac{c}{d}$ *and if* b *and* d *are both positive, then*

$$\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}.$$

Proof. If follows from $ad < bc$ that

$$a(b+d) = ab + ad < ab + bc = b(a+c)$$

Therefore

$$\frac{a}{b} < \frac{a+c}{b+d}.$$

Similarly

$$\frac{a+c}{b+d} < \frac{c}{d}.$$

∎

5.5.4. Theorem. $(1 + a_1)(1 + a_2) \ldots (1 + a_n) > 1 + (a_1 + a_2 + \ldots + a_n)$ *if* $n \geqslant 2$ *and* $a_i > 0$.

Proof. $(1 + a_1)(1 + a_2) \ldots (1 + a_n) = 1 + (a_1 + a_2 + \ldots + a_n) + C$ where

$C > 0$ since $n \geqslant 2$ and $a_i > 0$. By Rule 5.3.3,

$$1 + (a_1 + a_2 + \ldots + a_n) + C > 1 + (a_1 + a_2 + \ldots + a_n)$$

Therefore $(1 + a_1)(1 + a_2) \ldots (1 + a_n) > 1 + (a_1 + a_2 + \ldots + a_n)$. ∎

5.5.5. Theorem. $(1 - b_1)(1 - b_2) \ldots (1 - b_n) > 1 - (b_1 + b_2 + \ldots + b_n)$ *if* $n \geqslant 2$ *and* $0 < b_i < 1$ *for* $i = 1, 2, \ldots, n$.

Proof. The theorem is to be proved by induction on n.

For $n = 2$, we have $(1 - b_1)(1 - b_2) = 1 - (b_1 + b_2) + b_1 b_2$.
Since $b_i > 0$, $b_1 b_2 > 0$. Therefore $(1 - b_1)(1 - b_2) > 1 - (b_1 + b_2)$.
Suppose that the inequality

$$(1 - b_1)(1 - b_2) \ldots (1 - b_k) > 1 - (b_1 + b_2 + \ldots + b_k)$$

holds for some $k \geqslant 2$. Then by Rule 5.3.4,

$$(1 - b_1)(1 - b_2) \ldots (1 - b_k)(1 - b_{k+1})$$
$$> [1 - (b_1 + b_2 + \ldots + b_k)](1 - b_{k+1})$$
$$= 1 - (b_1 + b_2 + \ldots + b_k + b_{k+1}) + (b_1 + b_2 + \ldots + b_k) b_{k+1}$$

since $1 - b_{k+1} > 0$. Also $(b_1 + b_2 + \ldots + b_k) b_{k+1} > 0$ since $b_i > 0$.
Therefore by Rules 5.3.2 and 5.3.3,

$$(1 - b_1)(1 - b_2) \ldots (1 - b_{k+1}) > 1 - (b_1 + b_2 + \ldots + b_{k+1}).$$

Hence the theorem holds for all $n \geqslant 2$ by induction. ∎

5.5.6. Theorem. *For* $n = 1, 2, 3, \ldots$

$$a_1 a_2 \ldots a_n \leqslant \left(\frac{a_1 + a_2 + \ldots + a_n}{n} \right)^n$$

if $a_i \geqslant 0$ $(i = 1, 2, \ldots, n)$.

This has been proved in Example 2.7.2 by induction.

5.5.7. Theorem. *Let* $a_1, a_2, \ldots, a_n$ *and* $b_1, b_2, \ldots, b_n$ *be 2n real numbers. Then the Cauchy-Schwarz inequality*

$$(a_1 b_1 + a_2 b_2 + \ldots + a_n b_n)^2 \leqslant (a_1^2 + a_2^2 + \ldots + a_n^2)(b_1^2 + b_2^2 + \ldots b_n^2)$$

holds. Moreover, it holds with equality if and only if there are constants k and h, not both zero, such that $ka_i = hb_i$ *for all* $i = 1, 2, \ldots, n$ *(i.e.* a_i *and* b_i *are proportional).*

Proof. The proof of the inequality is based on a well-known property of quadratic equation that

$$Ax^2 + 2Bx + C = 0$$

has (i) two distinct real roots, (ii) one single real root, or (iii) no real root if the discriminant $4B^2 - 4AC$ is respectively (i) greater than zero, (ii) equal to zero, or (iii) less than zero.

Put

$$A = a_1^2 + a_2^2 + \ldots + a_n^2$$
$$B = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$
$$C = b_1^2 + b_2^2 + \ldots + b_n^2$$

and consider the quadratic equation

$$Ax^2 + 2Bx + C = 0.$$

Upon substitution, the left-hand side of the equation becomes

$$\begin{aligned} Ax^2 + 2Bx + C &= (a_1^2 + a_2^2 + \ldots + a_n^2)x^2 + 2(a_1 b_1 + a_2 b_2 + \ldots + a_n b_n)x + \\ &\quad (b_1^2 + b_2^2 + \ldots + b_n^2) \\ &= (a_1 x + b_1)^2 + (a_2 x + b_2)^2 + \ldots + (a_n x + b_n)^2 \end{aligned}$$

which is always non-negative for any real value of x. Therefore the quadratic equation $Ax^2 + 2Bx + C = 0$ can never have two distinct real roots. It follows that the discriminant $4B^2 - 4AC \leqslant 0$. Thus

$$B^2 \leqslant AC,$$

proving the Cauchy-Schwarz inequality.

For the second statement of the theorem concerning the equality $B^2 = AC$, we may assume that not all the a_i are zero, otherwise there is nothing to prove.

Suppose that there exist k and h such that $ka_i = hb_i$. Since not all the a_i are zero, $h \neq 0$. Therefore,

$$B = \frac{k}{h} A \quad \text{and} \quad C = \frac{k^2}{h^2} A.$$

Thus $B^2 = AC$.

Conversely suppose that $B^2 = AC$. Then the equation $Ax^2 + 2Bx + C = 0$ has a single real root $x = -\dfrac{B}{A}$. By substitution, we get

$$0 = (-\frac{a_1 B}{A} + b_1)^2 + (-\frac{a_2 B}{A} + b_2)^2 + \ldots + (-\frac{a_n B}{A} + b_n)^2.$$

Hence for $i = 1, 2, \ldots, n$

$$\frac{a_i B}{A} = b_i$$

proving that the a_i and b_i are proportional. ∎

Finally we remark that the notion of absolute value of a real number is entirely similar to that of an integer. Thus for any real number a, *the absolute value* of a is given by

$$|a| = \begin{cases} a \text{ if } a \geqslant 0 \\ -a \text{ if } a < 0 \end{cases}$$

$$|a| = \sqrt{a^2}$$

or $$|a| = \max(a, -a).$$

Moreover the following theorem can be proved in the same manner as its counterpart in Theorem 4.1.1.

5.5.8. Theorem. *For any real numbers a and b the following statements hold*

(a) $-|a| \leqslant a \leqslant |a|$

(b) $|ab| = |a||b|$

(c) *Let c be a positive real number.* $|a| < c$ *if and only if* $-c < a < c$.

(d) $|a + b| \leqslant |a| + |b|$ *(triangle inequality)*

(e) $|a - b| \geqslant |a| - |b|$.

5.6. EXERCISE

1. Prove that for $a, b \geqslant 0$ and any positive integer n,
$$a^n + b^n \leqslant (a + b)^n.$$

2. If $a + b + c = 2$, prove that $ab + bc \leqslant 1$.

3. If a, b and c are positive real numbers, prove that
$$(a + b)(b + c)(c + a) \geqslant 8abc.$$

4. Show that $a^2b + b^2c + c^2a + ab^2 + bc^2 + ca^2 \geqslant 6abc$ for all non-negative real numbers a, b and c.

5. Show that $a^2 + b^2 + c^2 \geqslant ab + bc + ca$ for any real numbers a, b and c.

6. If $a \geqslant b$ and $c \geqslant d$, prove that
$$\frac{ac + bd}{2} \leqslant \frac{a + b}{2} \cdot \frac{c + d}{2}.$$

7. If a, b, c and d are real numbers such that
$$a^2 + b^2 = 1 \quad \text{and} \quad c^2 + d^2 = 1,$$
prove that $ac + bd \leqslant 1$.

8. If a, b and c are real numbers such that $a + b + c = 1$, prove that $ab + bc + ca \leqslant 1/2$.

9. Let $a_1, a_2, \ldots, a_n$ be n positive real numbers. Their arithmetic mean A_n and harmonic mean H_n are defined by
$$A_n = \frac{\Sigma_{i=1}^{n} a_i}{n} \quad \text{and} \quad \frac{1}{H_n} = \frac{1}{n} \Sigma_{i=1}^{n} \frac{1}{a_i}.$$
Deduce from the Cauchy-Schwarz inequality that $H_n \leqslant A_n$.

*10. Let n be an integer greater than 1.
(a) If x is a real number greater than 1, show that
$$nx^{n-1}(x-1) > x^n - 1 > n(x-1).$$
(b) Hence prove that, if $0 < q < p$, then
$$np^{n-1}(p-q) > p^n - q^n > nq^{n-1}(p-q).$$
(c) Deduce from one of the inequalities in (b) that no positive real number has more than one positive n-th root.

*11. If x is a non-zero real number, show that $x + 1/x$ has no value between -2 and 2.

12. If a, b, c and d are positive real numbers, prove that
$$(a+b+c+d)^2 \geqslant 4(ab + bc + cd + da).$$

*13. If all the letters denote positive real numbers, show that
$$\frac{\Sigma_{i=1}^{n} a_i}{\Sigma_{i=1}^{n} b_i}$$
is not greater than the greatest and not less than the least of the fractions a_i/b_i $(i = 1, 2, 3, \ldots, n)$.

*14. (a) If the values $a_1, a_2, \ldots, a_k$ are all non-negative, and if $m_1, m_2, \ldots, m_k$ are positive integers, prove that
$$\frac{\Sigma_{i=1}^{k} m_i a_i}{M} \geqslant \left(\prod_{i=1}^{k} a_i^{m_i} \right)^{1/M}$$
where $M = \Sigma_{i=1}^{k} m_i$.

(b) Hence show that, if $q_1, q_2, \ldots, q_k$ are proper fractions satisfying $\sum_{i=1}^{k} q_k = 1$, then

$$\sum_{i=1}^{k} q_i a_i \geqslant \prod_{i=1}^{k} a_i^{q_i}.$$

15. If n is a positive integer and x is a positive real number not equal to 1, prove that

$$\frac{x^{n+1} - 1}{n+1} > \frac{x^n - 1}{n}.$$

16. If $a + b > 0$, prove that

$$\frac{a^3 + b^3}{2} \geqslant \left(\frac{a+b}{2}\right)^3.$$

17. $x_1, x_2, \ldots, x_n$ are n real numbers whose arithmetic mean is $\bar{x}$ and a is any real number other than $\bar{x}$. Prove that

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 < \sum_{i=1}^{n} (x_i - a)^2.$$

18. If a is a real number, prove that

$$\frac{a^4 - 1}{4} \geqslant \frac{a^3 - 1}{3}.$$

19. If a, b and c are non-zero real numbers, prove that

$$\frac{a}{b} + \frac{b}{c} + \frac{c}{a} \leqslant \frac{1}{2}\left(a^2 + b^2 + c^2 + \frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right).$$

For what values of a, b and c is the equality sign valid?

20. If a, b and c are three positive real numbers in ascending order of magnitude, show that

$$\frac{1}{a+c-b} < \frac{1}{a} + \frac{1}{c} - \frac{1}{b}.$$

21. If n is a positive integer and $x > 1$, deduce from Theorem 5.5.6 that

$$x^{2n+1} - 1 > (2n+1)(x-1)x^n.$$

22. If a, b and c are the lengths of the sides of a triangle, prove that

$$\frac{1}{a+b-c} + \frac{1}{b+c-a} + \frac{1}{c+a-b} > \frac{9}{a+b+c}.$$

23. If a, b and c are three real numbers and m is the smallest of the numbers $|a-b|$, $|b-c|$ and $|c-a|$. Show that
$$3(a^2+b^2+c^2) \geqslant (a-b)^2+(b-c)^2+(c-a)^2,$$
and hence that $a^2+b^2+c^2 \geqslant 2m^2$.

*24. By considering the sum $\Sigma_{i=1}^{n}(a_i-a_k)(b_i-b_k)$ or otherwise, show that if $\{a_1, a_2, \ldots, a_n\}$ and $\{b_1, b_2, \ldots, b_n\}$ are two sets of real numbers, arranged in descending order of magnitude, then
$$(\Sigma_{i=1}^{n} a_i)(\Sigma_{i=1}^{n} b_i) \leqslant n\,\Sigma_{i=1}^{n} a_i b_i \quad \text{(Tchebychef's inequality)}$$
holds and that the equality holds if and only if
$$a_1 = a_2 = \ldots = a_n \quad \text{or} \quad b_1 = b_2 = \ldots = b_n.$$

*25. If a, b and c are integers, not all zero, prove that
$$3(a^2+b^2+c^2) - 2(ab+bc+ca) \geqslant 3.$$

*26. If $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ are real numbers, prove that
$$\sqrt{\Sigma_{i=1}^{n}(a_i+b_i)^2} \leqslant \sqrt{\Sigma_{i=1}^{n} a_i^2} + \sqrt{\Sigma_{i=1}^{n} b_i^2} \quad \text{(Minkowski's inequality)}.$$

*27. If n is a positive integer greater than 1, by using Theorem 5.5.6 and by considering the sums
$$\Sigma_{i=1}^{n} i \quad \text{and} \quad \Sigma_{i=1}^{n} \frac{1}{i(i+1)}$$
show that
$$\left(\frac{n+1}{2}\right)^n > n! > (n+1)^{(n-1)/2}.$$

28. Sketch the graph of the function $f(x) = |x|$.

29. Prove that, for a given real number $\epsilon > 0$, $|x-a| < \epsilon$ if and only if
$$a-\epsilon < x < a+\epsilon.$$

30. For any two real numbers a and b, prove that
$$|a-b| \geqslant ||a|-|b||.$$

31. Prove that $a^2 < b^2$ if and only if $|a| < |b|$.

32. If a and b are real numbers such that $|a-b| < c$ for any positive real number c, prove that $a = b$.

33. For any real numbers a and b, prove that
    (a) $|a+b| \leqslant |a| + |b|$ (triangle inequality) and
    (b) $|a-b| \geqslant |a| - |b|$.

34. Show that $|a_1 + a_2 + \ldots + a_n| \leqslant |a_1| + |a_2| + \ldots + |a_n|$ for any real numbers $a_1, a_2, \ldots, a_n$.

35. Show that $\max(a, b) = \dfrac{a+b+|b-a|}{2}$

    and $\min(a, b) = \dfrac{a+b-|b-a|}{2}$

5.7. DENSENESS OF THE RATIONAL NUMBERS

The first six sections of this chapter deal with the fundamental algebraic properties of the real number system $\mathbb{R}$. These are all consequences of our basic assumptions or rules on the arithmetic and the order relation of real numbers. A careful reading of these rules will reveal that they are also valid for the smaller system $\mathbb{Q}$ of rational numbers. Thus all the results that we have obtained so far hold for the system $\mathbb{R}$ as well as for the system $\mathbb{Q}$. In other words, we have not yet touched on the most characteristic properties, namely the geometric properties, of the system $\mathbb{R}$. In this section we shall discuss one such property, the denseness of rational numbers in $\mathbb{R}$, which can be derived from Rule 5.3.6, the Archimedean postulate.

In Section 5.1, we have seen that the set $\mathbb{Q}$ of all rational numbers is a proper subset of the set $\mathbb{R}$ of all real numbers, i.e. there are irrational numbers in $\mathbb{R}$. Geometrically the system $\mathbb{R}$ is represented by a straight line, the number line $\mathbb{R}$. Thus on the number line we shall find two kinds of points, the rational points and the irrational points, representing the rational numbers and the irrational numbers respectively. We now wish to study the pattern in which these two kinds of points are distributed along the number line.

Let us begin by choosing two points on a straight line to represent the numbers 0 and 1, say 1 on the right of 0. Then with the segment between 0 and 1 as the unit length, we may mark off the positive integers on the right-hand side of 0 and the negative integers on the left-hand side of 0.

Then we bisect the unit length and use a segment of length 1/2 to mark off multiples of 1/2. Similarly, after trisecting the unit length, we may mark off multiples of 1/3, and so on (Fig. 5.6).

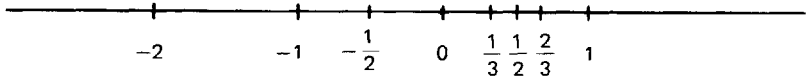


Fig. 5.6

Naturally it is impossible for us to complete the task of marking off all rational points on the number line, because there are an infinite number of them, but the position of each given rational number can be located in this manner, one at a time. In any case, we can imagine that rational points can be found 'everywhere' along the number line as confirmed by the following *denseness theorem.*

5.7.1. Theorem. *Between any two real numbers there is one rational number.*

Proof. Let a and b be two real numbers with $a < b$. An application of the Archimedean postulate to the positive numbers $b - a$ and 1 would yield a positive integer n such that $1 < n(b - a)$ and $n > 0$.

Hence $$\frac{1}{n} < b - a$$

We claim that at least one of the multiples

$$\ldots, -\frac{3}{n}, -\frac{2}{n}, -\frac{1}{n}, 0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \ldots$$

of $1/n$ must lie between a and b. Suppose that this were not the case. Then we would be able to find two consecutive multiples

$$\frac{x}{n} \quad \text{and} \quad \frac{x+1}{n}$$

of $1/n$ with both a and b lying between them,

i.e. $$\frac{x}{n} \leqslant a < b \leqslant \frac{x+1}{n}.$$

It would then follow that

$$\frac{1}{n} = \frac{x+1}{n} - \frac{x}{n} \geqslant b - a > \frac{1}{n}$$

which is absurd. Therefore at least one of the multiples, a rational number m/n, must lie between a and b, i.e. $a < m/n < b$. ∎

5.7.2. Corollary. *Between any two real numbers there are infinitely many rational numbers.*

Proof. Let $a < b$ be two real numbers. Then by Theorem 5.7.1, we can find a rational number x_0 such that

$$a < x_0 < b, \quad x_0 \in \mathbb{Q}.$$

Apply Theorem 5.7.1 to $x_0 < b$ we find x_1 such that

$$x_0 < x_1 < b, \quad x_1 \in \mathbb{Q}.$$

Using the same argument repeatedly, we can find for each natural number $n \in \mathbb{N}$, a rational number $x_n \in \mathbb{Q}$ such that

$$a < x_0 < x_1 < x_2 < \ldots < x_n < x_{n+1} < \ldots < b.$$

The detail of the induction is left as an exercise. ∎

We may describe the above property of the distribution of the rational numbers by saying that *the set of rational numbers is everywhere dense.* This means that in every interval of the number line, no matter how small, there are rational numbers. Visually the rational numbers appear to be so dense along the number line that there would be no room left for any more points. However the existence of irrational numbers such as $\sqrt{2}$ (see Example 5.1.1) proves beyond any doubt that even after all the rational numbers have taken their place in the number line, there are still plenty of 'holes' left for the irrational numbers. Using sophisticated methods of set theory, one can actually prove that in a way there are far more irrational points than rational points on the number line.

5.8. POSTULATE OF CONTINUITY

In this section we shall formulate our last basic assumption on the real numbers, the postulate of continuity. If we imagine the number line as being first populated by the integers and then by the rational numbers, then this postulate has something to do with filling the vacant positions on the number line so that the real numbers (rational and irrational together)

would form a continuum without interruption. We shall go into this aspect of the postulate in the subsequent sections. However, the full significance of the postulate will only become apparent in the next chapter, where we make use of it to study the important notions of limit and convergence, which are fundamental to calculus.

Let us consider a few subsets of $\mathbb{R}$:

(1) All prime numbers.
(2) All integers less than 10.
(3) All positive integers less than 10.
(4) All integers which are prefect squares.
(5) All real numbers x such that $-1 \leqslant x \leqslant 3$.
(6) All real numbers x such that $-3 < x < -2$.
(7) All real numbers x such that $0 \leqslant x \leqslant 5$ or $-4 \leqslant x \leqslant -1$.

Except (3), these sets are all infinite subsets of $\mathbb{R}$. When represented on the number line, (5) and (6) do not show up any gaps in them while all the others do. These two are of a type which we call *finite intervals.* There is, however, a difference between them: (5) contains its end points -1 and 3, and is called a *closed interval;* (6) does not contain its end points -3 and -2, and is called an *open interval*. We use the notation $[a, b]$ to denote a closed interval; thus $[a, b] = \{x \in \mathbb{R} | a \leqslant x \leqslant b\}$. Similarly we use the notation (a, b) to denote an open interval; thus $(a, b) = \{x \in \mathbb{R} | a < x < b\}$. Moreover we take note that (5) and (6) are finite intervals but they are both infinite sets. On the number line this is indicated as in Fig. 5.7 below:

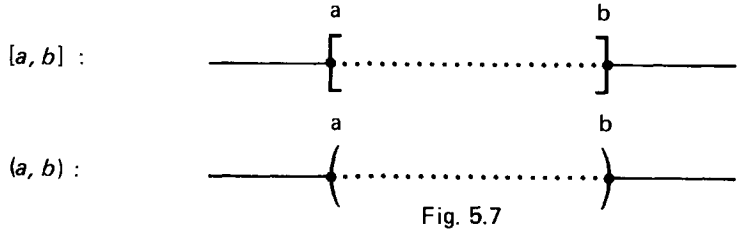


Fig. 5.7

Of the seven subsets of $\mathbb{R}$ above, some have a greatest element and/or a least element while others do not. The *greatest element* of a subset S of $\mathbb{R}$ is defined as the real number u which satisfies the two conditions:

$$\text{(i) } u \in S \quad \text{and} \quad \text{(ii) } u \geqslant x \text{ for all } x \in S,$$

and is often written as max S. Similarly the *least element* of S is defined as

the real number v that satisfies

$$\text{(iii) } v \in S \quad \text{and} \quad \text{(iv) } v \leqslant x \text{ for all } x \in S,$$

and is noted by $\min S$. If S is a finite non-empty subset, then S has both a greatest element and a least element. For example, the finite set (3) has the numbers 1 and 9 as its least element and greatest element respectively. If S is an infinite set, then it may or may not have such elements. For example, the closed interval $[a, b]$ has $a = \min[a, b]$ and $b = \max[a, b]$; but the interval $T = \{x \in \mathbb{R} \mid 0 \leqslant x < 1\}$, which is neither closed nor open, has only a least element $\min T = 0$, while $\max T$ does not exist. Because for any $x \in T$ we can always find a $y \in T$ such that $x < y$ (for instance, $y = (1 + x)/2$ satisfies $x < y < 1$).

Let us consider the definitions of something weaker by dropping the conditions (i) and (iii).

5.8.1. Definition. *Let S be a set of real numbers. If there is a real number K such that*

$$x \leqslant K \text{ for every } x \in S,$$

we say that S is bounded above. *K is called an* upper bound *of S. Similarly if there is a k such that*

$$k \leqslant x \text{ for every } x \in S,$$

then S is bounded below *and k is a* lower bound *of S. If S is bounded both above and below, we may simply say that it is* bounded. *A set which is not bounded is called* unbounded.

The three different cases may be illustrated by Fig. 5.8 below.

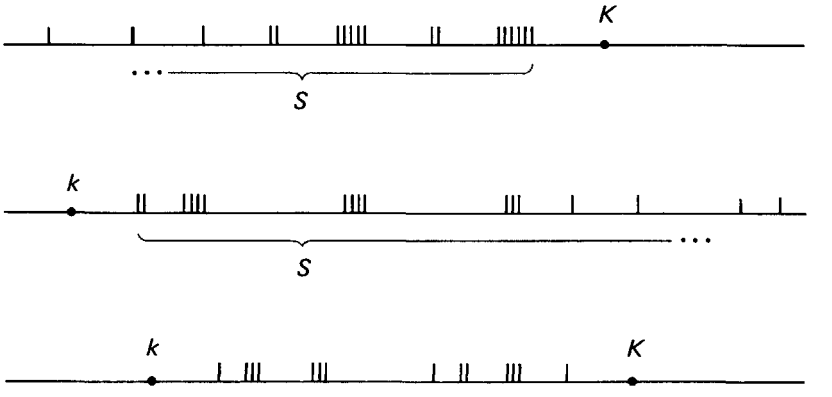


Fig. 5.8

If a set S has a greatest element max S, then S is bounded above and max S is an upper bound of S. Similarly for the least element. In particular every finite set of real numbers has a greatest element and a least element; therefore it is bounded.

The set $S = \{n/(n+1) \mid n \in \mathbb{Z}\} = \{0, 1/2, 2/3, 3/4, \ldots\}$ has $\min S = 0$. Therefore it is bounded below. S does not have a greatest element because for every $n/(n+1)$ belonging to S, we have $(n+1)/(n+2)$ also belonging to S such that $n/(n+1) < (n+1)/(n+2)$. But S is bounded above because any real number not less than 1 is an upper bound of S.

If S has an upper bound K, then any real number greater than K is also an upper bound. Since every upper bound of S can serve as a kind of barrier over which S cannot extend, the best barrier is obviously the *least upper bound*, which is a real number K such that

(a) K is an upper bound of S, i.e. $x \leqslant K$ for all $x \in S$;

(b) $K - \epsilon$ (where ϵ is any positive number however small) would not be an upper bound of S, i.e. for every $\epsilon > 0$ there is an $x \in S$ such that $K - \epsilon < x$.

In other words, the least upper bound of S is an upper bound K of S such that any number strictly less than K is no more an upper bound of S. Similarly we seek the *greatest lower bound* as the best barrier on the left.

It is easy for us to recognize the similiarity of these notions with the notions of greatest common divisor and least common multiple of number theory. If we recall that the existence of the greatest common divisor and the least common multiple depends on the well-ordering principle, we shall not be surprised that existence of the greatest lower bound and the least upper bound also depends on some important assumption. It is precisely for this reason that we should have the postulate of continuity to serve as a foundation stone of mathematical analysis.

5.8.2. Postulate of continuity. *Every non-empty set of real numbers, which is bounded above, has a least upper bound. Every non-empty set of real numbers, which is bounded below, has a greatest lower bound.*

5.8.3. *Example.* The set $S = \{n/(n+1) \mid n \in \mathbb{Z}\}$ is bounded. Therefore by the postulate of continuity, S has a greatest lower bound and a least upper bound. The greatest lower bound is 0 which is identical with the least element $\min S$. The least upper bound is 1. We have seen earlier that 1 is

an upper bound of S. Let $\epsilon > 0$ be any positive number. We may choose a positive integer n such that $n > (1 - \epsilon)/\epsilon$. Then $1 - \epsilon - \epsilon n < 0$. It follows that $1 - \epsilon - \epsilon n + n < n$ and $(1 - \epsilon)(n + 1) < n$. Hence $1 - \epsilon < n/(n + 1)$, i.e. $1 - \epsilon$ is no more an upper bound of S. Therefore 1 is the least upper bound of S.

5.8.4. *Example.* The set $T = \{x \in \mathbb{Q} | x^2 < 2\}$ is obviously bounded above by 2. Therefore it has a least upper bound. This is the real number $\sqrt{2}$. $\sqrt{2}$ is obviously an upper bound. To see that it is the least of the upper bounds, we may apply the denseness Theorem 5.7.1. to the interval $(\sqrt{2} - \epsilon, \sqrt{2})$, where ϵ is a positive number, to find a rational number x such that $\sqrt{2} - \epsilon < x < \sqrt{2}$.

Because of their importance, we find it convenient to call the least upper bound and the greatest lower bound of a set the *supremum* and the *infimum* respectively.

5.8.5. Definition. *Let S be a set of real numbers. The* supremum *of S is a real number K such that*

(a) $x \leqslant K$ for every $x \in S$, and

(b) for every positive ϵ there is an $x \in S$ for which $K - \epsilon < x$.

In this case we write $K = \sup S$. The infimum *of S is a real number k such that*

(c) $k \leqslant x$ for every $x \in S$, and

(b) for every positive ϵ there is an $x \in S$ for which $x < k + \epsilon$.

In this case we write $k = \inf S$.

5.9. EXERCISE

1. Let S be the set of all rational numbers r such that $0 < r < 1$. Show that S has no maximum and no minimum.

2. Suppose a non-empty set S of real numbers is bounded above and $T \subset S$. Prove that $\sup T \leqslant \sup S$.

3. Let f be the function defined as follows:

$$f(x) = \begin{cases} x & \text{if } x \text{ is a rational number,} \\ 0 & \text{if } x \text{ is an irrational number} \end{cases}$$

and let $g(x) = \sup\{f(x') : x' \leqslant x\}$. Prove that

$$g(x) = \begin{cases} x & \text{if } x \geqslant 0, \\ 0 & \text{if } x < 0. \end{cases}$$

4. Suppose S is the set of all rational numbers less than a certain real number a. Assuming that between any two real numbers there is a rational number, prove that $\sup S = a$.

5. Prove that the set $S = \{\frac{n-1}{n} : n \text{ is a positive integer}\}$ is bounded above with $\sup S = 1$. Does S have a maximum?

6. Suppose S is a non-empty set of real numbers which is bounded above and $T = \{-x : x \in S\}$, prove that $\inf T = -\sup S$.

7. Suppose a non-empty set S of real numbers is bounded above and a is a given real number. If $T = \{x + a : x \in S\}$, prove that
$$\sup T = \sup S + a.$$

8. Suppose a non-empty set S of real numbers is bounded above and $a > 0$. If $T = \{ax : x \in S\}$, prove that
$$\sup T = a \sup S.$$

9. Let c be a positive real number.
   (a) If n is any integer greater than $1/c$, show that there is a positive integer m such that
   $$\frac{m^2}{n^2} < c \leqslant \frac{(m+1)^2}{n^2}.$$
   Hence deduce that $c - \frac{m^2}{n^2} \leqslant \frac{3m}{n^2}$ and that $(\frac{3m}{n^2})^2 < \frac{9c}{n^2}$.
   (b) Let S be the set of all numbers which are squares of rational numbers and which are less than c. Prove that $c = \sup S$.

10. By Example 5.3.10, given any positive real number c, there is a decimal approximation $a_0.\, a_1 a_2 \ldots a_n$ of c such that
$$a_0.\, a_1 a_2 \ldots a_n \leqslant c < a_0.\, a_1 a_2 \ldots a_n + 10^{-n}.$$
Denote by A the set of all such decimal approximations of c. Does $\sup A$ exist? What is $\sup A$?

*11. Given two positive real numbers a and b, consider the set

$$T = \{t \mid t \in \mathbb{N} \text{ and } ta \leqslant b\}.$$

(a) Prove that T is non-empty and bounded above and hence has a supremum (i.e. a least upper bound).

(b) Let $n = 1 + \sup T$. Prove that $na > b$. Does it mean that the Archimedean postulate is a consequence of the postulate of continuity?

*12. Let S be a non-empty set of natural numbers. As a set of real numbers, S is bounded below (e.g. by -2). Therefore by the principle of continuity S has an infimum (i.e. a greatest lower bound).

(a) Prove that $\inf S \in S$.

(b) Prove that $\inf S = \min S$.

Does this mean that the well-ordering principle as well as the principle of mathematical induction are consequences of the postulate of continuity?

*13. Let a be a number greater than 1.

(a) Let S be a set of positive real numbers. Prove that $\inf S > 0$ if and only if there is a number x in S such that $x < a \inf S$.

(b) Let R be the set of all numbers a^{-n}, where $n = 1, 2, 3, \ldots$. Prove that $\inf R = 0$.

*14. The distance $d(a, S)$ between a real number a and a non-empty set S of real numbers is defined by

$$d(a, S) = \inf T.$$

where $T = \{|a - x| : x \in S\}$.

(a) If $a \in S$, prove that $d(a, S) = 0$.

(b) If S is bounded above and $a = \sup S$, prove that $d(a, S) = 0$. Deduce that the same is true if S is bounded below and $a = \inf S$.

5.10. POWERS AND ROOTS

Let a be a real number and k a positive integer. Then we denote by a^k the product of k identical factors a. Thus

$$a^k = aa \ldots a \ (k \text{ factors}).$$

a^k is called the *k-th power* of the number a. We sometimes call a the *base*

and k the *index* or *exponent* of the power a^k. If $a \neq 0$, we follow the convention in putting

$$a^0 = 1 \text{ and } a^{-k} = \frac{1}{a^k} \quad (a \neq 0).$$

The simple rules,

$$a^k\, a^m = a^{k+m}$$
$$a^k\, b^k = (ab)^k$$
$$(a^k)^m = a^{km}$$

are easily verified. Similarly it is not difficult to show that the following rules on inequalities hold.

5.10.1. Theorem. *Let k be a positive integer. For positive real numbers a and b,*

(a) $a^k < b^k$ *if and only if* $a < b$

(b) $a^k < a^{k+1}$ *if and only if* $1 < a$

(c) $a^{k+1} < a^k$ *if and only if* $a < 1$.

Taking roots of a positive number is the operation inverse to the forming of powers. We have seen earlier that the square root $\sqrt{2}$ of the rational number 2 fails to be a rational number. Thus the operation of taking roots cannot always be carried out successfully within the confine of the number system $\mathbb{Q}$. In this respect, we may say that the system is incomplete. In the next section we shall see that the extended number system $\mathbb{R}$ does not have this deficiency and that roots of positive real numbers are again real numbers. Meanwhile, we shall assume that our system $\mathbb{R}$ has this property and study roots of positive real numbers in some detail.

Let us first review the definition of roots. Given a non-negative real number a and a positive integer k, we say that the non-negative real number b is a *k-th root* of a if

$$b^k = a.$$

In this case, we write

$$b = \sqrt[k]{a}.$$

In other words, $b = \sqrt[k]{a}$ is defined by the two conditions:

$$b^k = a \quad \text{and} \quad b \geqslant 0$$

Thus $\sqrt[2]{9} = 3$ and $\sqrt[2]{9} \neq -3$ although $(-3)^2 = 9$. For square roots where $k = 2$, we usually omit the superscript k; thus $\sqrt[2]{a} = \sqrt{a}$ and $\sqrt[2]{2} = \sqrt{2}$.

It follows from the rules for powers that the following rules on roots hold.

For non-negative real numbers a and b,

$$\sqrt[k]{a}\sqrt[k]{b} = \sqrt[k]{ab}$$
$$(\sqrt[k]{a})^m = \sqrt[k]{a^m}$$
$$\sqrt[k]{\sqrt[m]{a}} = \sqrt[km]{a}.$$

Similarly the following rules on inequalities hold.

5.10.2. Theorem. *For positive real numbers a and b,*

(a) $\sqrt[k]{a} < \sqrt[k]{b}$ *if and only if* $a < b$
(b) $\sqrt[k+1]{a} < \sqrt[k]{a}$ *if and only if* $1 < a$
(c) $\sqrt[k]{a} < \sqrt[k+1]{a}$ *if and only if* $a < 1$.

Proof. (a) follows immediately from Theorem 5.10.1 (a).

(b) If $1 < a$, then by Theorem 5.10.1(b), $a^k < a^{k+1}$. Taking the $k(k+1)$-th root on both sides, we get

$$\sqrt[k(k+1)]{a^k} < \sqrt[k(k+1)]{a^{k+1}}$$
$$\therefore \sqrt[k+1]{\sqrt[k]{a^k}} < \sqrt[k]{\sqrt[k+1]{a^{k+1}}}$$
$$\therefore \sqrt[k+1]{a} < \sqrt[k]{a}$$

Conversely if $\sqrt[k+1]{a} < \sqrt[k]{a}$, then taking the $k(k+1)$-th power on both sides, we get $a^k < a^{k+1}$. Therefore, $1 < a$ by Theorem 5.10.1(b).

(c) can be proved in a similar manner. ∎

Another notation for the k-th root $\sqrt[k]{a}$ of a non-negative real number a is the exponential form $a^{1/k}$. Thus for any $a \geqslant 0$ and any positive integer k, $b = a^{1/k}$ is the real number such that

$$b^k = a \quad \text{and} \quad b \geqslant 0$$

From this it is easy to extend our definition of power to any rational exponents. If $a > 0$ and $r = m/k$ is a rational number, where $k > 0$, then we define

$$a^r = (a^m)^{1/k}.$$

Then it follows from the two sets of rules that the following rules of rational exponents hold.

5.10.3. Theorem. *For any positive number a and b and rational exponents r and s,*

$$a^r a^s = a^{r+s}$$
$$a^r b^r = (ab)^r$$
$$(a^r)^s = a^{rs}$$
$$\frac{a^r}{a^s} = a^{r-s}.$$

Combining Theorems 5.10.1 and 5.10.2, we have

5.10.4 Theorem. *For positive real numbers a and b and positive rational exponents $r < s$,*

(a) $a^r < b^r$ *if and only if* $a < b$

(b) $a^r < a^s$ *if and only if* $1 < a$

(c) $a^s < a^r$ *if and only if* $a < 1$.

5.11. EXISTENCE OF ROOTS

Though the subject of our discussion in this section and the next is of some importance to a rigorous study of the real numbers, we can only offer here the sketch of a process by which the roots of a positive real number can be obtained. There are two reasons for the brevity. First, a full description of the process would be beyond the scope of the present course. Second, some of the missing details will be supplied in the next chapter. If the reader finds the presentation natural and comprehensible, so much the better. If he finds it difficult to follow, he need not be disheartened, but should continue with the subsequent sections and only return to it after reading the next chapter.

The tool for finding the k-th root is the method of decimal approximation as described in Example 5.3.10. Recall that given a positive real number c, its n-th decimal approximation is given by the decimal number $a_0 . a_1 a_2 \ldots a_n$ such that for $i = 0, 1, 2, \ldots, n$

$$a_0 . a_1 a_2 \ldots a_i \leqslant c < a_0 . a_1 a_2 \ldots a_i + 10^{-i}$$

where a_0 is a natural number and the decimal digits $a_1, a_2, \ldots, a_n$ have values 0, 1, 2, . . . , 8 or 9.

Let a be a positive real number and k a positive integer. We shall follow the method of decimal approximation to find an increasing sequence of

decimal numbers $b_n = x_0 . x_1 x_2 \ldots x_n$ $(n = 0, 1, 2, \ldots)$ such that

$$(x_0 . x_1 x_2 \ldots x_n)^k \leqslant a < (x_0 . x_1 x_2 \ldots x_n + 10^{-n})^k$$

The number x_0 and the digits $x_1, x_2, \ldots, x_n, \ldots$ are to be found recursively.

To begin with, x_0 is the least natural number such that

$$a < (x_0 + 1)^k .$$

Clearly x_0 exists by the well-ordering principle.

Suppose that the number x_0 and the digits $x_1, x_2, \ldots, x_n$ are found and satisfy the required condition. Then we divide the very small interval $[x_0 . x_1 \ldots x_n, x_0 . x_1 \ldots x_n + 10^{-n}]$ into 10 equal subintervals (Fig. 5.9).

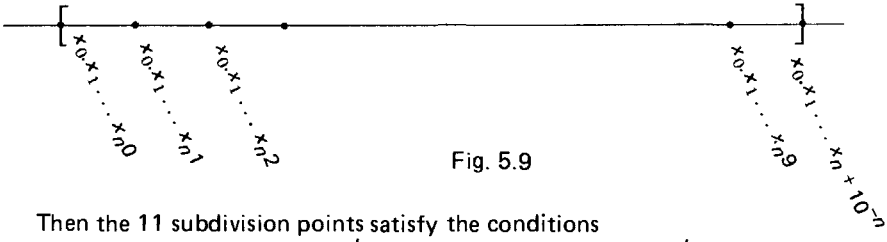


Fig. 5.9

Then the 11 subdivision points satisfy the conditions

$$(x_0 . x_1 \ldots x_n 0)^k \leqslant a < (x_0 . x_1 \ldots x_n + 10^{-n})^k$$

and $(x_0 . x_1 \ldots x_n 0)^k < (x_0 . x_1 \ldots x_n 1)^k < (x_0 . x_1 \ldots x_n 2)^k < \ldots$

$$< (x_0 . x_1 \ldots x_n 9)^k < (x_0 . x_1 \ldots x_n + 10^{-n})^k .$$

Therefore the next digit x_{n+1} can be found among 0, 1, 2, . . . , 9 such that

$$(x_0 . x_1 \ldots x_{n+1})^k \leqslant a < (x_0 . x_1 \ldots x_{n+1} + 10^{-(n+1)})^k .$$

Hence, by induction we have, for $n = 0, 1, 2, \ldots$

$$b_n = x_0 . x_1 \ldots x_n \quad \text{and} \quad c_n = x_0 . x_1 \ldots x_n + 10^{-n}$$

such that

$$0 \leqslant b_0 \leqslant b_1 \leqslant \ldots \leqslant b_n \leqslant b_{n+1} \leqslant \ldots < c_n < \ldots < c_1 < c_0$$

and

$$0 \leqslant b_0^k \leqslant b_1^k \leqslant \ldots \leqslant b_n^k \leqslant b_{n+1}^k \leqslant \ldots \leqslant a < \ldots < c_n^k < \ldots < c_1^k < c_0^k .$$

Consider the sets of decimal numbers

$$B = \{b_0, b_1, b_2, \ldots\} \quad \text{and} \quad C = \{c_0, c_1, c_2, \ldots\} .$$

Both sets are bounded. Therefore by the postulate of continuity, they have greatest lower bounds and least upper bounds. It follows that

$$\sup B = \inf C \quad \text{and} \quad (\sup B)^k = a = (\inf C)^k .$$

Therefore

$$\sup B = \sqrt[k]{a} .$$

If we may use the terminology of the next chapter, the above statements can be rephrased as: the sequences $\langle b_n \rangle$ and $\langle c_n \rangle$ both converge to the same limit $b = \sqrt[k]{a}$.

5.12. POWERS AND LOGARITHM

We have seen in section 5.10 that given a positive real number a and a rational number r, the power a^r with rational exponent is defined as the real number

$$a^r = (a^m)^{1/k}$$

where $r = m/k$ with $m, k \in \mathbb{Z}$ and $k > 0$. Moreover, for the calculation of such powers, we have the rules listed in Theorems 5.10.3 and 5.10.4.

We can use the method of decimal approximation of Example 5.3.10 to extend our definition of power to arbitrary real exponents.

Let $a > 1$ be a real number and $c > 0$ be a positive real number. Then c can be approximated by a sequence of decimal numbers

$$c_n = x_0 . x_1 x_2 \ldots x_n$$

such that $0 \leqslant c_n \leqslant c < c_n + 10^{-n}$,

$$0 \leqslant c_0 \leqslant c_1 \leqslant \ldots \leqslant c_n \leqslant \ldots \leqslant c$$

and $c < \ldots < c_n + 10^{-n} < \ldots < c_1 + 10^{-1} < c_0 + 1.$

Since all the c_n are rational numbers, by Theorem 5.10.4 we have

$$0 < a^{c_0} \leqslant a^{c_1} \leqslant \ldots \leqslant a^{c_n} \leqslant a^{c_{n+1}} \leqslant \ldots$$

The set $S = \{a^{c_0}, a^{c_1}, \ldots, a^{c_n}, \ldots\}$ is bounded below by 0 and bounded above by a^{c_0+1}. Therefore by the postulate of continuity, S has a least upper bound sup S, which is a real number. We define the power a^c with base a and exponent c as the real number

$$a^c = \sup S$$

where $a > 1$ and $c > 0$.

For the other cases, we put

$$a^c = \frac{1}{a^{-c}} \qquad \text{where } a > 1, c < 0$$

$$a^c = \left(\frac{1}{a}\right)^{-c} \qquad \text{where } 0 < a < 1, c \neq 0$$

and $a^0 = 1$ where $a > 0$.

Thus for any positive real number a and arbitrary real number c, the power

a^c is a well-defined real number. Obviously this general definition of power includes the ordinary power a^k and root $\sqrt[k]{a}$ with positive integer k as special cases. For the calculation of such general powers, we have rules similar to those listed in Theorems 5.10.3 and 5.10.4.

5.12.1. Theorem. *For any positive real numbers a and b and arbitrary exponents c and d,*

(a) $a^c a^d = a^{c+d}$

(b) $a^c b^c = (ab)^c$

(c) $(a^c)^d = a^{cd}$

(d) $\dfrac{a^c}{a^d} = a^{c-d}$

5.12.2. Theorem. *For any positive real numbers a and b and positive exponents* $0 < c < d$,

(a) $a^c < b^c$ *if and only if* $a < b$

(b) $a^c < a^d$ *if and only if* $1 < a$

(c) $a^d < a^c$ *if and only if* $a < 1$.

Unlike their counterparts Theorems 5.10.3 and 5.10.4, the proof of the two theorems above is no easy task. We do not wish to hold up our progress with further attempts in establishing these rules by approximation, but ask our reader to accept them as valid.

Finally we want to conclude this chapter with a brief description on the notion of logarithm. Again, we shall make use of the method of decimal approximation. Given real numbers $b > 1$ and $x > 0$, the logarithm $\log_b x$ of x with base b is the real number c such that

$$b^c = x.$$

The real number c can be obtained by a sequence of decimal approximations $c_n = h_0 . h_1 h_2 \ldots h_n$ such that

$$b^{c_n} \leqslant x < b^{c_n + 10^{-n}}$$

for all $n = 0, 1, 2, \ldots$. In other words, the sequence of integers $h_0, h_1, h_2, \ldots$ are to be found inductively. The integer h_0 in front of the decimal point is the integer such that

$$b^{h_0} \leqslant x < b^{h_0 + 1}$$

Then by subdividing the interval $[h_0, h_0 + 1]$, we find the next digit h_1 such that

$$b^{h_0 . h_1} \leqslant x < b^{h_0 . h_1 + 10^{-1}}.$$

Further subdivision of the interval $[h_0 . h_1, h_0 . h_1 + 10^{-1}]$ yields the next digit h_2, and so forth. The sequence of increasing decimal numbers

$$c_0 \leqslant c_1 \leqslant c_2 \leqslant \ldots \leqslant c_n \leqslant \ldots$$

is bounded above, and hence has a least upper bounde $\sup\{c_0, c_1, \ldots\} = c$ such that

$$b^c = x.$$

Now we put

$$c = \log_b x.$$

The rules for calculation with logarithm follow immediately from Theorems 5.12.1 and 5.12.2. These are listed below.

5.12.3. Theorem. *Let $b > 1$ and let x and y be positive real numbers. Then the following statements holds.*

(a) $\log_b 1 = 0$ *and* $\log_b b = 1$

(b) $\log_b (xy) = \log_b x + \log_b y$

(c) $\log_b (\frac{x}{y}) = \log_b x - \log_b y$

(d) $\log_b x^d = d \log_b x$

(e) $\log_b x < \log_b y$ *if and only if* $x < y$

(f) $\log_b x > 0$ *if and only if* $x > 1$

(g) $\log_b x = \log_b c \log_c x$ $(c > 1)$.

Finally we take note that $\log_b a$ is only defined when the base b is greater than 1 and a is a positive real number. In all other cases, the logarithm is not defined. Thus the logarithm of a negative number or zero is meaningless.

5.13. EXERCISE

1. Solve the simultaneous equations

$$\begin{cases} 3^{x+3} = 9^{2-y}, \\ (\frac{1}{3})^y = 9^{x-4}. \end{cases}$$

2. Solve the equation

$$\log_{10} 250 - \log_{10} x = \frac{\log_{10} 64}{\log_{10} 4}$$

3. If a and b are unequal positive real numbers, prove that $a^a\, b^b > a^b\, b^a$.

4. Show that $\log_2 12$ is irrational.

5. Find the value of $\log_5 32 \cdot \log_4 7 \cdot \log_{49} 125$.

6. Show that $(\log_{10} 2)(2 \log_4 5 + 1) = 1$.

7. Solve the simultaneous equations
$$\begin{cases} xy = 16, \\ \log_x y = 3. \end{cases}$$

8. Solve the equation $\log_3 x = \log_9 (x + 6)$.

9. Solve the equation $\log_2 x + \log_x 2 = 2$.

10. If a, m and n are positive real numbers such that $a \neq 1$ and $m > n$, prove that
$$a^m + \frac{1}{a^m} > a^n + \frac{1}{a^n}\ .$$

11. Solve simultaneously $\begin{cases} x = 16y, \\ \log_y x - \log_x y = \dfrac{8}{3}. \end{cases}$

6. *Limit and Convergence*

Traditionally mathematics is divided into three large categories – geometry, algebra and analysis – though at times it is very difficult to put some modern topics of mathematics in one of these categories. A very rough outline of the classification can be given as follows. The study of space inaugurated by the ancient civilizations of the world, and summarized by Euclid in his *Elements,* together with the variations, generalizations and associated studies that have since been created, is called geometry. The abstract symbolic study of ordinary arithmetic is called algebra. Again, it includes a multitude of variations, generalizations and associated studies of the same subject. Finally analysis consists of those branches of mathematics that are allied to or arise from the calculus. It includes calculus, differential equations, integration theory, etc. The idea of a limit plays an important role in each of these studies. Generally, the presence of the concept of limit distinguishes analysis from algebra. It is precisely this important notion of limit that we shall study in this chapter.

6.1. NULL SEQUENCE

An *infinite sequence* (or simply a *sequence*) is a list of numbers $a_1, a_2, a_3, \ldots, a_n, \ldots$. The number a_n is called the n-th *term* of the sequence. The notation $\langle a_n \rangle$ represents the sequence whose n-th term is a_n ($n = 1, 2, \ldots$). Thus $\langle 1 \rangle = 1, 1, \ldots$ is the sequence whose n-th term is 1 for all $n = 1, 2, \ldots$, and $\langle (-1)^n \rangle = -1, 1, -1, 1, \ldots$ is the sequence where n-th term is $(-1)^n$.

By a *neighbourhood* of a point a on the number line $\mathbb{R}$, we mean an open interval which contains the point a. For example, $U = (-1, 1) = \{x : -1 < x < 1\}$ is a neighbourhood of 0. Since U also contains the point $-\frac{1}{2}$, it is also a neighbourhood of $-\frac{1}{2}$. In fact U is a neighbourhood of every point that it contains. Given any positive number ϵ (read 'epsilon') and any point a on the number line, the open interval

$$V = (a - \epsilon, a + \epsilon) = \{x : a - \epsilon < x < a + \epsilon\} = \{x : |x - a| < \epsilon\}$$

is an interval of length 2ϵ with mid-point a. V is a neighbourhood of a and is called an *ϵ-neighbourhood* of a.

One type of sequences that possess a particular property in relation to the number 0 (zero or null) shall play a very special role in our discussion. They are called the *null sequences.* Before we formulate a definition of null sequence, let us consider an example of such sequence:

$$\langle a_n \rangle = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$$

whose n-th term is the number $a_n = 1/n$.

We observe that all terms of the sequence $\langle a_n \rangle$ are positive and none of them is 0. Moreover, the n-th term $a_n = 1/n$ gets smaller as the index n gets larger. Thus $\langle a_n \rangle$ has the property that *the term a_n gets closer and closer to 0 as n increases.*

However this very descriptive property still does not bring out the special relationship between the sequence $\langle a_n \rangle$ and the number 0 because as n increases, a_n also gets closer and closer to $-1/2$, -1 as well as any negative number. A sharper observation would be that *the term a_n can get as close as we like to the number* 0.

The sequence $\langle a_n \rangle$ has indeed this property. For example, if we wish to have a term a_m within one millionth of the unit length from the number 0, we may take any term a_m with $m > 10^6$. This is an improvement upon our first observation because a_n cannot get within a distance of 10^{-1} to the number $-1/2$ nor within the same distance to the number -1. However, this description also suits other sequences which are of a different type. Take, for example, the sequence $\langle b_n \rangle$

$$\langle b_n \rangle = 1, \frac{1}{2}, \frac{1}{3}, \frac{3}{4}, \frac{1}{5}, \frac{5}{6}, \ldots$$

whose n-th term b_n is given by

$$b_n = \begin{cases} \dfrac{1}{n} & \text{if } n \text{ is odd} \\ 1 - \dfrac{1}{n} & \text{if } n \text{ is even.} \end{cases}$$

The term $b_{1000001}$ is within a distance of one millionth of the unit length from 0 while $b_{1000002}$ is within the same distance from 1, and terms still closer to 0 or to 1 can be found easily. The graphs of the sequences $\langle a_n \rangle$ and $\langle b_n \rangle$ will reveal their relationship to the number 0 more prominently.

From Figs. 6.1 & 6.2, we see that $\langle a_n \rangle$ has the property that *given any*

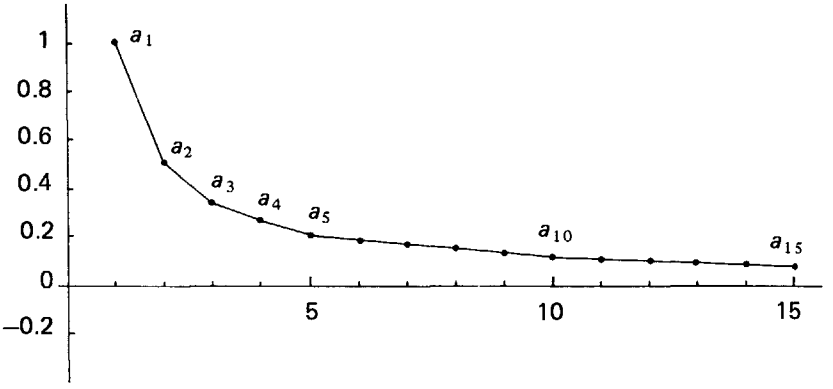


Fig. 6.1

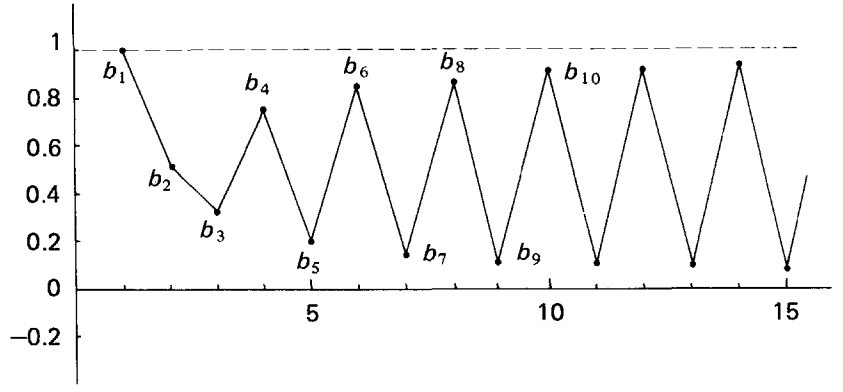


Fig. 6.2

neighbourhood U of 0, all the terms of the sequence $\langle a_n \rangle$, except a finite number of them (some initial segment), fall into the neighbourhood U, while $\langle b_n \rangle$ does not have this property.

For example, if $U = (-\epsilon, \epsilon)$ is an ϵ-neighbourhood of 0 with a very small $\epsilon = 10^{-6}$, then $a_n \in U$ for all $n > 10^6$, while the terms that do not fall into

U are those of the initial segment $a_1, a_2, \ldots, a_{1000000}$. On the other hand, the sequence $\langle b_n \rangle$ is seen to oscillate between the greatest lower bound 0 and the least upper bound 1; therefore $\langle b_n \rangle$ does not have the said property. Reformulating the above property in terms of distance from the point 0, we have the following formal definition of null sequence.

6.1.1. Definition. *A sequence $\langle a_n \rangle$ is called a null sequence if the following condition is satisfied: Given any positive $\epsilon > 0$ (no matter how small), there is a number N such that $|a_n - 0| < \epsilon$ for all $n > N$. In this case we also say that the sequence $\langle a_n \rangle$ converges or tends to 0 or that 0 is the limit of the sequence $\langle a_n \rangle$. The usual notation for this is* $\lim_{n\to\infty} a_n = 0$, *or* $\lim a_n = 0$, *or simply* $a_n \to 0$

6.1.2. *Remarks.* We take note that the number N in the above definition depends on the given positive number ϵ. Take $\langle a_n \rangle = 1, 1/2, 1/3, \ldots, 1/n, \ldots$ again. For $\epsilon = 1$, we find $N = 1$, or any other value will do. For $\epsilon = 10^{-1}$, N has to be at least 10. In general, we need larger values of N for smaller values of ϵ. Clearly $\langle a_n \rangle = 1, 1/2, 1/3, \ldots, 1/n, \ldots$ is a null sequence. If $\langle a_n \rangle$ is a null sequence and if ϵ and N form a pair that has the required property of Definition 6.1.1, then the infinite 'tail' $a_{N+1}, a_{N+2}, \ldots$ of the sequence $\langle a_n \rangle$ will fall in the neighbourhood $(-\epsilon, \epsilon)$ of 0, while possibly some of the terms among the finite 'initial segment' $a_1, a_2, \ldots, a_N$ may fail to do so. Therefore, being a null sequence is essentially a property of the 'tail' of a sequence $\langle a_n \rangle$ while the behaviour of any one 'initial segment' is unimportant.

6.1.3. *Example.* The *constant sequence* $\langle a_n \rangle = 0, 0, \ldots$ (i.e. $a_n = 0$ for all n) is clearly a null sequence because every term of the sequence is in any ϵ-neighbourhood of 0.

6.1.4. *Example* The sequence

$$\langle a_n \rangle = -1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, \ldots$$

whose n-th term is $a_n = \dfrac{(-1)^n}{n}$, is a null sequence. Thus $\lim_{n\to\infty} \dfrac{(-1)^n}{n} = 0$.

Notice that while $\langle 1/n \rangle$ approaches its limit 0 from the right-hand side

of the number line, the sequence $\langle (-1)^n/n \rangle$ tends to 0 oscillating about its limit 0.

6.1.5. *Example* Consider the sequence

$$\langle a_n \rangle = 1, \frac{1}{2}, \frac{1}{3}, 1, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{2}, \ldots$$

whose n-th term is given by

$$a_n = \begin{cases} \dfrac{1}{n} & \text{if } n \text{ is not a multiple of 4} \\ \dfrac{1}{m} & \text{if } n = 4m \text{ is a multiple of 4} \end{cases}$$

$\langle a_n \rangle$ is a null sequence. Thus $a_n \to 0$. In contrast to the previous examples, where the terms get steadily closer to the limit 0, it is not true that every term of the present sequence is closer to 0 than its preceding term.

6.1.6. *Example* Show that

$$\lim_{n\to\infty} \frac{n+9}{n^2+3n+7} = 0$$

Proof. Let $a_n = \dfrac{n+9}{n^2+3n+7}$ and $b_n = \dfrac{2}{n}$ for all $n = 1, 2, \ldots$ Then for all $n > 9$, we have

$$a_n = \frac{n+9}{n^2+3n+7} < \frac{2n}{n^2} = \frac{2}{n} = b_n.$$

These inequalities show that the sequence $\langle a_n \rangle$ is 'sandwiched' between the constant sequence $\langle 0 \rangle$ and the null-sequence $\langle b_n \rangle$ from $n = 10$ onwards. (Fig. 6.3).

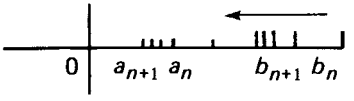


Fig. 6.3

As n increases, b_n approaches to its limit 0 from the right, pushing a_n, which is always to the left of b_n, towards 0 at the same time. Therefore $a_n \to 0$. ∎

6.2. EXERCISE

1. Find the limit of the sequence $\{u_n\}$ as $n \to \infty$ if $u_n = (n^2+4)/(n^3-2)$.

2. Suppose $\lim_{n\to\infty} a_n = 0$ and each $a_n > 0$. Prove that the set of all numbers a_n actually has a maximum member.

3. Show that (a) $\lim_{n\to\infty} \frac{1}{n^\alpha} = 0 \quad (\alpha > 0)$,
   (b) $\lim_{n\to\infty} (\sqrt{n+1} - \sqrt{n}) = 0$.

4. If, for some real number k and any positive integer n,
   $$0 < a_n < kb_n \quad \text{and} \quad \lim_{n\to\infty} b_n = 0,$$
   prove that $\lim_{n\to\infty} a_n = 0$.

5. If $|a| < 1$, prove that $a^n \to 0$ as $n \to \infty$.

6. Show that $\lim_{n\to\infty} (\sqrt[8]{n^2+1} - \sqrt[4]{n+1}) = 0$.

7. Evaluate $\lim_{n\to\infty} (\sqrt{n^2+1} - n)$.

8. If $\lim_{n\to\infty} a_n = 0$, show that $\lim_{n\to\infty} \lambda a_n = 0$ for any real number λ.

6.3. CONVERGENT SEQUENCE

To obtain a definition of convergent sequence we merely have to formulate the relationship between the terms of a proposed sequence $\langle a_n \rangle$ to a fixed number a similar to that between a null sequence and the number 0.

6.3.1. *Example.* Consider the sequence

$$\langle a_n \rangle = 2, \frac{1}{2}, \frac{4}{3}, \frac{3}{4}, \ldots, 1 + \frac{(-1)^{n-1}}{n}, \ldots$$

whose n-th term is $a_n = 1 + (-1)^{n-1}/n$. Some terms of the sequence are shown in Fig. 6.4 below.

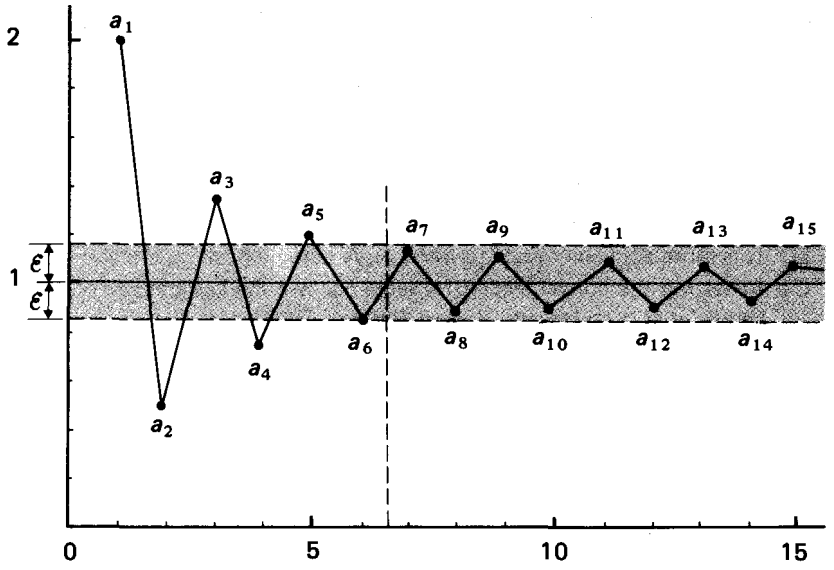


Fig. 6.4

Similar to the way in which the terms of a null sequence tend to the number 0 as n increases, the terms of the present sequence $\langle a_n \rangle$ tend to the number 1, in the sense that given any ϵ-neighbourhood of 1, one entire 'tail' of the sequence is trapped in the neighbourhood ($a_7, a_8, a_9, \ldots$ in Fig. 6.4). More precisely, given $\epsilon > 0$, let $N = 1/\epsilon$. Then

$$-\epsilon < (-1)^{n-1}/n < \epsilon$$

for all index $n > N$. Therefore $1 - \epsilon < a_n < 1 + \epsilon$ for all $n > N$, i.e. all the terms a_n of the sequence $\langle a_n \rangle$ with $n > N$ belong to the ϵ-neighbourhood $(1 - \epsilon, 1 + \epsilon)$ of 1. Thus we see that the relationship between a null sequence and the number 0 is entirely similar to the relationship between sequence $\langle a_n \rangle$ and the number 1. This leads us to say that the sequence $\langle a_n \rangle$ converges to the limit 1, and the following formal definition.

6.3.2. Definition. *A sequence $\langle a_n \rangle$ is said to converge to the limit a if given any $\epsilon > 0$, we can find a number N, such that $|a_n - a| < \epsilon$ for any $n > N$. We then write* $\lim_{n\to\infty} a_n = a$, $\lim a_n = a$ *or* $a_n \to a$.

6.3.3. *Remarks.*

(1) The number N in the above definition depends on the given positive number ϵ. To emphasize this dependence, some authors use the functional notation $N(\epsilon)$ instead of N. However, such notation may convey a wrong impression that the number N would be unique for each given ϵ which, as we have seen in the case of null sequences, is not true at all. Therefore the simple notation N is preferred.

(2) It is more convenient not to insist on N being a positive integer.

(3) In testing the convergence of a sequence, we shall emphasize on small values of $\epsilon > 0$. Usually the smaller the value of ϵ, the larger shall be the value of N. The following table shows some corresponding pair of values of ϵ and N for the sequence $\langle a_n \rangle$ of Example 6.3.1.

ϵ	N
0.8	1.25 or greater
0.45	2.22 or greater
0.1	10 or greater
0.01	100 or greater
10^{-10}	10^{10} or greater

(4) The symbol $n \to \infty$ just expresses the unending growth of n and is read 'as n tends to infinity'. We emphasize that infinity (∞) is not a number and the word 'infinity' has to follow 'tends to' in order to have a meaning.

(5) When a given sequence $\langle a_n \rangle$ is represented graphically as in Example 6.3.1, the statement $a_n \to a$ means that, however small ϵ is, there is a vertical line $x = N$ such that all represented points of the sequence to the right of it lie within the band between $y = a - \epsilon$ and $y = a + \epsilon$.

(6) $\lim a_n = a$ is clearly equivalent to $\langle a_n - a \rangle$ being a null sequence.

(7) A sequence that converges to a limit is called a *convergent* sequence.

(8) It is easily seen that the behaviour of a sequence as regards to convergence is unchanged if we omit or alter a finite number of terms a_n.

(9) Clearly all null sequences are convergent sequences.

6.3.4. *Example.* Let $\langle x_n \rangle$ be the sequence whose n-th term $x_n = d_0.d_1 \ldots d_n$ is the n-th decimal approximation of $\sqrt{2}$ such that

$$x_n^2 \leqslant 2 < (x_n + 10^{-n})^2.$$

Then $\lim_{n\to\infty} x_n = \sqrt{2}$.

6.3.5. *Example.* If a is a real number, then the constant sequence $\langle a_n \rangle = a, a, \ldots, a, \ldots$ ($a_n = a$ for all n) tends to a.

6.3.6. *Example.* If a is a positive real number, then $\lim \sqrt[n]{a} = 1$.

Proof. We have three cases to consider according to the value of a: $a = 1$, $a > 1$ or $a < 1$. The first case where $a = 1$ is trivial.

Suppose that $a > 1$. Then by Theorem 5.10.2, we have $\sqrt[n]{a} > 1$. Therefore we can write

$$\sqrt[n]{a} = 1 + h_n$$

where h_n is a positive real number. By Theorem 5.5.4 we get

$$a = (1 + h_n)^n > 1 + nh_n.$$

It follows that

$$0 < h_n < \frac{a-1}{n}$$

and

$$\sqrt[n]{a} = 1 + h_n < 1 + \frac{a-1}{n}.$$

This means that the distance between $\sqrt[n]{a}$ and 1 is less than $(a-1)/n$. As n increases this distance tends to 0, which means that $\sqrt[n]{a}$ tends to 1.

Finally, suppose $a < 1$, then $\sqrt[n]{a} < 1$. If we write

$$\sqrt[n]{a} = \frac{1}{1 + k_n}$$

where k_n is a positive real number, then

$$a = \frac{1}{(1+k_n)^n} < \frac{1}{1+nk_n}$$

and

$$0 < k_n < \frac{1}{n}\left(\frac{1}{a} - 1\right).$$

As n increases, k_n tends to 0 and hence $\sqrt[n]{a} = 1/(1 + k_n)$ tends to 1. Therefore, we have proved that in all cases $\sqrt[n]{a} \to 1$ for $a > 0$. ∎

6.4. DIVERGENT SEQUENCE

Sequences that do not converge (to a limit) are called *divergent sequences.* Among the divergent sequences, some have terms that grow beyond all bounds.

6.4.1. *Example.* The sequence $\langle a_n \rangle = a, 2a, 3a, \ldots$ whose n-th term is $a_n = na$, where a is a positive number, is a divergent sequence. Let x be a real number and $\epsilon > 0$ a positive number. Then by the Archimedean postulate, for every number N there is a positive integer n such that $N < n$ and $x + \epsilon < na$, i.e. $|a_n - x| > \epsilon$. Therefore $\langle a_n \rangle$ does not converge to any number x. In other words, the sequence $\langle a_n \rangle$ is divergent because its terms a_n grow indefinitely as n increases. To express this state of affair, we write

$$a_n \to \infty$$

(read a_n tends to infinity). Again the symbol ∞ does not represent any real number and the notation '$a_n \longrightarrow \infty$' means that a_n is greater than any given number x for all sufficiently large n.

Similarly we have divergent sequences whose terms diminish beyond all bounds.

6.4.2. *Example.* The sequence $\langle b_n \rangle = -1, -\sqrt{2}, -\sqrt{3}, \ldots$ whose n-th term is $b_n = -\sqrt{n}$ is a divergent sequence. Clearly b_n can be chosen to be less than any given $x - \epsilon$, i.e. $|b_n - x| > \epsilon$. In this case we write

$$b_n \to -\infty$$

(read b_n tends to negative infinity). Like ∞, the symbol $-\infty$ does not represent a real number. In particular, it makes no sense at all to write

$$\infty + a, \quad -\infty + b, \quad \infty + \infty, \quad \infty - \infty, \text{ etc.}$$

There are also divergent sequences that neither tend to infinity nor to negative infinity.

6.4.3. *Example.* The sequence $\langle c_n \rangle = 1, 1/2, 1/3, 3/4, 1/5, 5/6,$ whose n-th term is given by

$$c_n = \begin{cases} \dfrac{1}{n} & \text{if } n \text{ is odd} \\ 1 - \dfrac{1}{n} & \text{if } n \text{ is even} \end{cases}$$

is a divergent sequence. $\langle c_n \rangle$ does not converge to any number but

oscillates between 0 and 1, for as n increases, it gets closer and closer to 0 or 1, according to whether n is odd or even. Therefore it is divergent and tends neither to infinity nor to negative infinity.

6.5. EXERCISE

1. Find $\lim_{n\to\infty} \dfrac{3n^2+1}{n^2-5n}$.

2. If $\lim_{n\to\infty} a_n = l$, prove that $\lim_{n\to\infty} |a_n| = |l|$.

3. (a) If $0 < a < 2$, prove that $a < \sqrt{2a} < 2$.
   (b) Prove that the sequence $\sqrt{2}, \sqrt{2\sqrt{2}}, \sqrt{2\sqrt{2\sqrt{2}}}, \ldots$ converges.
   (c) Find the limit of the sequence in (b).

4. If $x_{n+2} = (x_{n+1} + x_n)/2$, $(n = 0, 1, 2, \ldots)$, prove that
$$\lim_{n\to\infty} x_n = (x_0 + 2x_1)/3.$$

5. Let $x_{n+1} = \dfrac{2a^2 x_n}{x_n^2 + a^2}$ and $y_{n+1} = \dfrac{y_n^2 + a^2}{2y_n}$ $(n = 0, 1, 2, \ldots)$
   If $x_0 > 0$, $y_0 > 0$ and $a > 0$, by considering $(a - x_{n+1})/(a + x_{n+1})$ and $(y_{n+1} - a)/(y_{n+1} + a)$, prove that $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = a$.

*6. (a) If $y = (ax + b)/(x + c)$ and λ, μ are the roots of the equation
$$x^2 + (c - a)x - b = 0,$$
   show that
$$\frac{y-\lambda}{y-\mu} = \left(\frac{c+\mu}{c+\lambda}\right)\frac{x-\lambda}{x-\mu}.$$
   (b) If $a_n a_{n+1} - p(a_n - a_{n+1}) = q^2$ for all natural numbers n and $pa_0 > 0$, prove that
$$\lim_{n\to\infty} a_n = \begin{cases} q & \text{if } \dfrac{p}{q} > 0, \\ -q & \text{if } \dfrac{p}{q} < 0, \\ 0 & \text{if } q = 0. \end{cases}$$

7. If $a > 1$, prove that $a^n \to \infty$ as $n \to \infty$.

8. Let q be a positive rational number. Prove that $n^q \to \infty$ as $n \to \infty$.

6.6. SUM, PRODUCT AND QUOTIENT OF CONVERGENT SEQUENCES

In the previous sections, it has not been too difficult to guess the limits of the convergent sequences even if they are not explicitly given. In general, it is not easy to determine whether a given sequence converges and what its limit is. There is no standard computational procedure that always gives the correct answer about limit and convergence, but there are some useful methods which are applicable to a variety of sequences. The first of these methods involves some algebraic operations on the terms of known convergent sequences.

Consider, for all $n = 1, 2, \ldots,$

$$b_n = 1 \quad \text{and} \quad c_n = (-1)^n/n.$$

Clearly the constant sequence $\langle b_n \rangle$ converges to 1, while $\langle c_n \rangle$ is a null sequence. Taking the sums of the corresponding terms of these two sequences:

$$a_n = b_n + c_n = 1 + (-1)^n/n$$

we obtain the terms of the convergent sequence $\langle a_n \rangle$ of Example 6.3.1, which is seen to tend to 1 as n increases. This suggests that we may obtain useful results by looking into some algebraic operations on convergent sequences.

6.6.1. Theorem. *Let $\langle a_n \rangle$ and $\langle b_n \rangle$ be two convergent sequences with $a_n \to a$ and $b_n \to b$. Then $\langle a_n + b_n \rangle$ and $\langle a_n b_n \rangle$ are convergent sequences such that*

$$\lim (a_n + b_n) = \lim a_n + \lim b_n = a + b$$
$$\lim (a_n\, b_n) = \lim a_n \cdot \lim b_n = ab$$

Moreover if $b_n \neq 0$ for all n and $b \neq 0$, then $\langle a_n/b_n \rangle$ is a convergent sequence such that

$$\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n} = \frac{a}{b}$$

Proof. For the convergence of the sum $\langle a_n + b_n \rangle$ we have to show that the sequence $\langle (a_n + b_n) - (a + b) \rangle$ is a null sequence. This leads us to consider the following inequality:

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leqslant |a_n - a| + |b_n - b|$$

The right-hand side and hence also the left-hand side of the above inequality can be made as small as we like for a sufficiently large n. Therefore the sequence in question is a null sequence. Alternatively, we can work according to the prescription of Definition 6.3.2. Given any $\epsilon > 0$, as $\epsilon/2 > 0$, we have N_1 and N_2 such that

$$|a_n - a| < \epsilon/2 \quad \text{for all } n > N_1$$

$$|b_n - b| < \epsilon/2 \quad \text{for all } n > N_2 .$$

Therefore for all $n > \max(N_1, N_2)$ we get

$$|(a_n + b_n) - (a + b)| \leqslant |a_n - a| + |b_n - b| < \epsilon .$$

Either way we have shown $\lim(a_n + b_n) = a + b$.

Similarly for the convergence of the product $\langle a_n\, b_n \rangle$, we may use the inequality:

$$\begin{aligned} |a_n b_n - ab| &= |(a_n - a)\, b_n + (b_n - b)\, a| \\ &\leqslant |a_n - a|\, |b_n| + |b_n - b||a| \end{aligned}$$

Now it follows from $b_n \to b$ that for the positive number $|b|$, there is an index N such that if $n > N$, then $|b_n - b| < |b|$ and hence $|b_n| < 2|b|$. Thus if we put

$$K = \max(|b_1|, |b_2|, \ldots, |b_N|, 2|b|)$$

then $|b_n| < K$ for all n. Therefore we obtain

$$|a_n b_n - ab| \leqslant |a_n - a|\, K + |b_n - b||a|$$

from which $a_n b_n \to ab$ follows.

Finally for the quotient, it is sufficient to consider the special case $\lim 1/b_n = 1/b$ where $b \neq 0$ and $b_n \neq 0$ for all n. These conditions on the limit and the terms of $\langle b_n \rangle$ imply that there is a positive number L such that $L \leqslant |b_n|$ for all n. To see this, we first find an index N such that if $n > N$, then $|b_n - b| < |b|/2$, and hence

$$|b_n| = |b + (b_n - b)| \geqslant |b| - |b_n - b| > |b|/2.$$

Thus we may put

$$L = \min(|b_1|, |b_2|, \ldots, |b_N|, |b|/2) .$$

Now the convergence of $\langle 1/b_n \rangle$ follows from the following inequality

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \left| \frac{b - b_n}{b_n\, b} \right| \leqslant \left| \frac{1}{Lb} \right| |b_n - b| \qquad \blacksquare$$

6.6.2. *Example.* How does

$$s_n = \frac{-n^2 + 5n + 7}{3n^2 - 2n - 6}$$

behave as $n \to \infty$?

Solution. Rewrite s_n as

$$s_n = \frac{(-n^2 + 5n + 7)/n^2}{(3n^2 - 2n - 6)/n^2} = \frac{-1 + \frac{5}{n} + \frac{7}{n^2}}{3 - \frac{2}{n} - \frac{6}{n^2}}.$$

Consider separately

$$a_n = -1 + \frac{5}{n} + \frac{7}{n^2} \text{ and } b_n = 3 - \frac{2}{n} - \frac{6}{n^2}.$$

Now $\langle a_n \rangle$ is the sum of the sequences $\langle -1 \rangle$, $\langle 5/n \rangle$ and $\langle 7/n^2 \rangle$. By Theorem 6.6.1, we get $\lim a_n = -1$. Similarly $\lim b_n = 3$. Therefore $\lim s_n = \lim (a_n/b_n) = -1/3$. ∎

It follows from this that it is entirely justified to approach the same problem in an intuitive manner by thinking that the terms $5n$ and 7 of the numerator of s_n will become negligible in comparison with the leading term $-n^2$ for very large values of n. Similarly, we may think that $-2n$ and -6 of the denominator of s_n will be insignificant in comparison with the leading term $3n^2$ for large values of n. Therefore as n increases indefinitely, s_n becomes about the same as the quotient $-n^2/3n^2 = -1/3$ of the leading terms, and hence $\lim s_n = -1/3$. In other words, the highest indices in the numerator and the denominator will predominate and determine the limit of the quotient. Thus as $n \longrightarrow \infty$,

$$\frac{n^2 + 1}{n^2 - 1} \longrightarrow 1$$

$$\frac{-n^2 + 5}{n^3 + 2n^2 + 1} \longrightarrow 0$$

$$\frac{n^3 + 2n^2 - 1}{100n^2 + 7n + 1} \longrightarrow \infty.$$

6.7. THE SANDWICH THEOREM

Another effective method for the evaluation of limit is the so-called 'sandwich' theorem, a special case of which has been used already in Example 6.1.6. The idea is to put the proposed sequence $\langle a_n \rangle$ between two convergent sequences $\langle x_n \rangle$ and $\langle y_n \rangle$ which tend to the same limit a. As n increases, the interval between x_n and y_n gets smaller and smaller while

retaining the term a_n within. Thus as both x_n and y_n tend to a, a_n is forced to tend to the same limit a.

6.7.1. Sandwich Theorem. *Let* $\langle x_n \rangle$, $\langle y_n \rangle$ *and* $\langle a_n \rangle$ *be sequences. If* $\lim x_n = \lim y_n = a$ *and* $x_n \leqslant a_n \leqslant y_n$ *for all* n, *then* $\langle a_n \rangle$ *is a convergent sequence and* $\lim a_n = a$.

Proof. Graphically the three sequences are related to one another as below (Fig. 6.5):

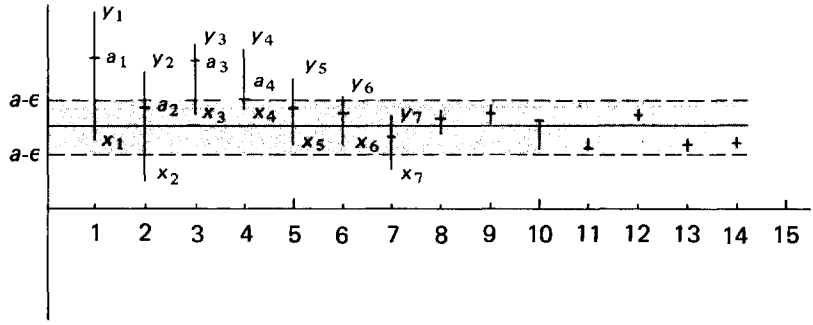


Fig. 6.5

The vertical bars are the intervals $[x_n, y_n]$ and the notches across are the values of a_n. The shaded horizontal band represents an ϵ-neighbourhood of a. Since both $x_n \to a$ and $y_n \to a$, the vertical bars will eventually (i.e. for sufficiently large n) fall within the shaded horizontal band. In terms of ϵ and N we proceed as follows:

Given any $\epsilon > 0$, let N_1 and N_2 be numbers such that

$$|x_n - a| < \epsilon \quad \text{for all } n > N_1$$
$$|y_n - a| < \epsilon \quad \text{for all } n > N_2.$$

Then for all $n > N = \max(N_1, N_2)$

$$a - \epsilon < x_n < a + \epsilon \quad \text{and} \quad a - \epsilon < y_n < a + \epsilon.$$

Since $x_n \leqslant a_n \leqslant y_n$ for all n,

$$a - \epsilon < x_n \leqslant a_n \leqslant y_n < a + \epsilon \quad \text{for all } n > N$$
$$\therefore \quad a - \epsilon < a_n < a + \epsilon$$
$$\therefore \quad |a_n - a| < \epsilon \quad \text{for all } n > N.$$

Hence $a_n \longrightarrow a$. ∎

A particularly useful corollary of the Sandwich Theorem is the following:

6.7.2. Corollary. *If $\langle y_n \rangle$ is a null sequence of positive terms and if $|a_n - a| \leqslant y_n$ for all $n = 1, 2, \ldots$, then the sequence $\langle a_n \rangle$ is convergent and* $\lim a_n = a$.

Proof. The inequality $|a_n - a| \leqslant y_n$ can be rewritten as

$$a - y_n \leqslant a_n \leqslant a + y_n .$$

By the sum rule, both $\langle a - y_n \rangle$ and $\langle a + y_n \rangle$ are convergent. Therefore $(a - y_n) \longrightarrow a$ and $(a + y_n) \longrightarrow a$ and hence $a_n \longrightarrow a$ by Theorem 6.7.1. ∎

6.7.3. *Example* If $0 < a < 1$ then $a^n \longrightarrow 0$ as $n \longrightarrow \infty$.

Proof. Put $a_n = a^n$ and $1/a = 1 + h$ where $h > 0$. Consider the sequences $\langle x_n \rangle$ and $\langle y_n \rangle$ whose n-th terms are respectively

$$x_n = 0 \quad \text{and} \quad y_n = \frac{1}{1 + nh}$$

Then it follows from

$$\frac{1}{a^n} = (1 + h)^n > 1 + nh \quad \text{for all } n \geqslant 2$$

that $0 < a_n < 1/(1 + nh)$. Therefore $x_n < a_n < y_n$ (for all $n \geqslant 2$). Since $x_n \longrightarrow 0$ and $y_n \longrightarrow 0$, it follows that $a_n \longrightarrow 0$, i.e. $a^n \longrightarrow 0$. ∎

6.7.4. *Example.* If $a_n = \sqrt{4n^2 - 5n + 2} - 2n$, then $a_n \longrightarrow \frac{5}{4}$ as $n \longrightarrow \infty$

Proof. By completing the square, we have

$$4n^2 + 5n + 2 = (2n + \frac{5}{4})^2 + \frac{7}{16} .$$

Then $$a_n - \frac{5}{4} = \sqrt{4n^2 + 5n + 2} - (2n + \frac{5}{4})$$

$$= \frac{(\sqrt{4n^2 + 5n + 2})^2 - (2n + \frac{5}{4})^2}{\sqrt{4n^2 + 5n + 2} + (2n + \frac{5}{4})}$$

$$= \frac{7/16}{\sqrt{4n^2 + 5n + 2} + (2n + \frac{5}{4})} < \frac{1}{2n} .$$

Therefore by Theorem 6.7.2, $\lim a_n = 5/4$. ∎

6.8. EXERCISE

1. Prove that $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

2. Let x be a positive real number and let N be the smallest integer greater than x. Prove that
$$\frac{x^n}{n!} \leqslant \frac{x^{N-1}}{(N-1)!}\left(\frac{x}{N}\right)^{n-N+1} \quad (n \geqslant N).$$
Deduce that $\frac{x^n}{n!} \to 0$ as $n \to \infty$.

3. Show that $\lim_{n\to\infty} \frac{n!}{n^n} = 0$.

4. Let a be any positive rational number and let $|x| < 1$.
   (a) Show that there exists a natural number N such that
   $$\left(1 + \frac{1}{N}\right)^{a+1} |x| \leqslant 1.$$
   (b) Deduce that $|n^{a+1} x^n| \leqslant |N^{a+1} x^N| \quad (n \geqslant N)$.
   (c) Hence show that $n^a x^n \to 0$ as $n \to \infty$.

5. (a) If $0 < a < b$, prove that
   $$a < \frac{2ab}{a+b} < \sqrt{ab} < b.$$
   (b) Two sequences $\{x_n\}$ and $\{y_n\}$ are defined inductively by
   $$x_1 = 1/2, \quad y_1 = 1,$$
   $$x_n = \sqrt{x_{n-1} y_{n-1}} \quad (n = 2, 3, 4, \ldots)$$
   and $$\frac{1}{y_n} = \frac{1}{2}\left(\frac{1}{x_n} + \frac{1}{y_{n-1}}\right) \quad (n = 2, 3, 4, \ldots).$$
   (i) Prove that $x_{n-1} < x_n < y_n < y_{n-1} \quad (n = 2, 3, 4, \ldots)$.
   (ii) Deduce that both sequences converge to the same limit l, where $1/2 < l < 1$. [Actually $l = \pi/4$.]

6. If $x_{n+1} = x_n^2 + 1/4$ for any positive integer n and $x_1 = a$, prove that $\{x_n\}$ is an increasing sequence. If $0 < a \leqslant 1/2$, prove that $\lim_{n\to\infty} x_n = 1/2$.
   If $a > 1/2$, prove that the sequence $\{x_n\}$ is unbounded.

7. A sequence $\{x_n\}$ is defined by the relation

$$x_1 = 1 \quad \text{and} \quad x_{n+1} = \sqrt{x_n + 1} \quad (n = 1, 2, 3, \ldots),$$

prove that $\lim_{n\to\infty} x_n = \dfrac{1+\sqrt{5}}{2}$.

8. Let $x > 0$. Prove that $\lim_{n\to\infty} x^{1/n} = 1$.

9. If $a > b > 0$ and a_n, b_n are defined by the recurrence relations

$$a_0 = a,$$
$$b_0 = b,$$
$$a_n = \frac{a_{n-1} + b_{n-1}}{2},$$
$$b_n = \frac{2a_{n-1}\, b_{n-1}}{a_{n-1} + b_{n-1}}.$$

(a) Prove that, for any positive integer n,

$$a > a_1 > a_2 > \ldots > a_n > b_n > \ldots > b_2 > b_1 > b.$$

and $\quad a_n - b_n < \dfrac{a-b}{2^n}$.

(b) Prove that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \sqrt{ab}$.

6.9. MONOTONE SEQUENCE

The methods of the last sections consist in comparing a proposed sequence, whose convergence is being tested with some known convergent sequences. Therefore, we should have a large reserve of known convergent sequences to give these methods a wide scope of applications. We shall see in this section that monotone sequences will provide us with a large supply of useful convergent sequences.

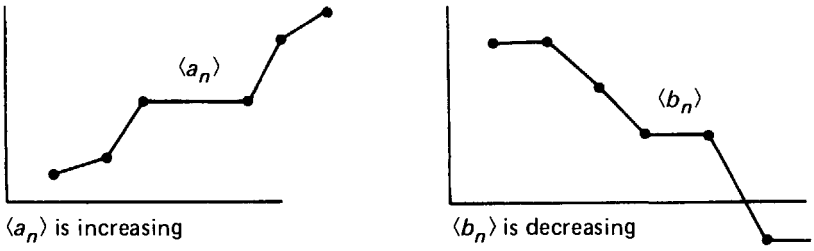
A sequence $\langle a_n \rangle$ is *increasing* (notation $a_n \nearrow$) if

$$a_n \leqslant a_{n+1} \quad (n = 1, 2, \ldots).$$

Similarly a sequence $\langle b_n \rangle$ is *decreasing* (notation $b_n \searrow$) if

$$b_n \geqslant b_{n+1} \quad (n = 1, 2, \ldots).$$

Increasing and decreasing sequences are called *montone sequences* (Fig. 6.6).



$\langle a_n \rangle$ is increasing $\langle b_n \rangle$ is decreasing

Fig. 6.6

We say that the sequence $\langle a_n \rangle$ is *strictly increasing* if $a_n < a_{n+1}$ $(n = 1, 2, \ldots)$ and that $\langle b_n \rangle$ is *strictly decreasing* if $b_n > b_{n+1}$ $(n = 1, 2, \ldots)$.

6.9.1. *Example.* $\langle (-1)^n \rangle$ is not a monotone sequence.

6.9.2. *Example.* If $a > 1$, then $\langle a^n \rangle$ is a strictly increasing sequence that tends to infinity.

6.9.3. *Example.* The constant sequence $\langle 1 \rangle = 1, 1, \ldots$ is an increasing (and also a decreasing) sequence that tends to 1 as $n \to \infty$.

6.9.4. *Example.* If $0 < b < 1$ then $\langle b^n \rangle$ is a strictly decreasing sequence that converges to 0.

6.9.5. *Example.* Let $\langle x_n \rangle$ be the sequence whose n-th term x_n is the n-th decimal approximation of $\sqrt{2}$. Then $\langle x_n \rangle$ is a increasing sequence that converges to $\sqrt{2}$.

We say that a sequence $\langle c_n \rangle$ is *bounded above* (respectively *bounded below*) if the set

$$\{c_n \mid n = 1, 2, \ldots\}$$

of its terms is bounded above (respectively bounded below). The monotone sequences of the last three examples all converge and satisfy the boundedness condition. More generally, we have the following important theorems. We shall see that in the proof of these theorems, the postulate of continuity will play a most crucial role.

6.9.6. Theorem. *An increasing sequence $\langle a_n \rangle$ converges if and only if it is bounded above. In this case as $n \to \infty$, a_n tends to the least upper bound (i.e. the supremum) of the set of terms of $\langle a_n \rangle$.*

6.9.7. Theorem *A decreasing sequence $\langle b_n \rangle$ converges if and only if it is bounded below. In this case as $n \to \infty$, b_n tends to the greatest lower bound (i.e. the infimum) of the set of terms of $\langle b_n \rangle$.*

Proof. Let $\langle a_n \rangle$ be an increasing sequence. If $\langle a_n \rangle$ is not bounded above, then $a_n \to \infty$ as $n \to \infty$. Therefore $\langle a_n \rangle$ is divergent.

Conversely assume that $\langle a_n \rangle$ is bounded above. By the postulate of continuity, the set $S = \{a_n \mid n = 1, 2, \ldots\}$ of terms of $\langle a_n \rangle$ has a least upper bound K which is a real number. We contend that $a_n \to K$ as $n \to \infty$ (Fig. 6.7).

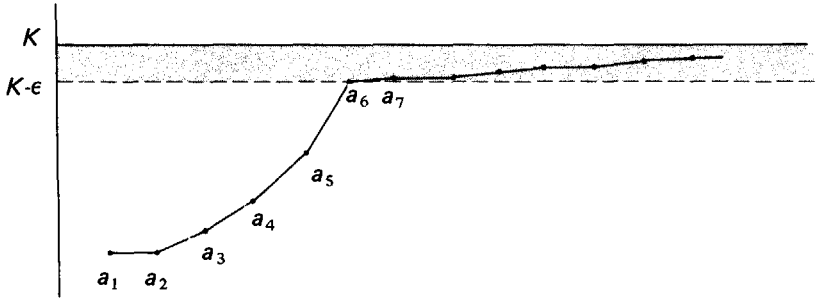


Fig. 6.7

By the definition of supremum, K has the properties

(i) $a_n \leqslant K$ for all $n = 1, 2, \ldots$

(ii) given any $\epsilon > 0$, $K - \epsilon < a_N$ for some N (i.e. $K - \epsilon$ is no more an upper bound of S).

Therefore for any ϵ-neighbourhood $(K - \epsilon, K + \epsilon)$ of K there is an index N such that $K - \epsilon < a_N$. Since $a_n \uparrow$, it follows that for $n > N$,

$$K - \epsilon < a_N \leqslant a_n.$$

On the other hand, by (i) above, $a_n \leqslant K$. Therefore

$$a_n < K + \epsilon.$$

Hence for all $n > N$,

$$K - \epsilon < a_n < K + \epsilon,$$

proving that $a_n \to K$ as $n \to \infty$.

The proof for the convergence of a decreasing sequence $\langle b_n \rangle$ which is bounded below is similar. Alternatively we may consider the increasing sequence $\langle -b_n \rangle$ and apply the result just obtained. ∎

6.9.8. *Example.* Let $x_n = 1 + 1/2 + 1/3 + \ldots + 1/n$ $(n = 1, 2, \ldots)$. Then the sequence $\langle x_n \rangle$ is a strictly increasing sequence. By Theorem 6.9.6, the sequence $\langle x_n \rangle$ is convergent or divergent according to whether it is bounded above or not. We shall see that it is not bounded above and hence divergent.

Let $G > 0$ be an arbitrary positive real number. By the Archimedean postulate, we can find a positive integer m such that $2G < m$. Now for any $n \geqslant 2^m$, we get

$$\begin{aligned} a_n &= 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} \geqslant 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{2^m} \\ &= (1 + \frac{1}{2}) + (\frac{1}{3} + \frac{1}{4}) + (\frac{1}{5} + \ldots + \frac{1}{8}) + \ldots + (\frac{1}{2^{m-1}+1} + \ldots + \frac{1}{2^m}) \\ &> \frac{1}{2} + 2 \times \frac{1}{4} + 4 \times \frac{1}{8} + \ldots + 2^{m-1} \times \frac{1}{2^m} \\ &= \frac{1}{2} + \frac{1}{2} + \ldots + \frac{1}{2} = \frac{m}{2} > G. \end{aligned}$$

Therefore a_n increases indefinitely without bound. Thus

$$(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}) \to \infty \text{ as } n \to \infty.$$

6.9.9. *Example.* Let $y_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$.

Then the sequence $\langle y_n \rangle$ is a strictly increasing sequence. We shall show below that $\langle y_n \rangle$ is bounded above and hence convergent.

It follows from the inequality $n! \geqslant 2^{n-1}$ that

$$\begin{aligned} y_n &= 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} \leqslant 1 + (1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}) \\ &= 1 + \frac{1 - (\frac{1}{2})^n}{1 - \frac{1}{2}} = 1 + 2\,(1 - (\tfrac{1}{2})^n) < 3. \end{aligned}$$

Thus $y_n < 3$ for all $n = 1, 2, \ldots$ Hence $\langle y_n \rangle$ is bounded above and convergent. We shall denote the limit of the sequence $\langle y_n \rangle$ by e, which is a very important constant in mathematics, and call the real number e the *base of natural logarithm.* Thus

$$\lim \left(1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}\right) = e.$$

Moreover since 3 is an upper bound of the sequence $\langle y_n \rangle$, it follow that $e \leqslant 3$ and

$$y_1 < y_2 < y_3 < \ldots < e \leqslant 3.$$

The terms of $\langle y_n \rangle$ can be taken as approximations of the constant e; the larger the index n, the better the approximation of e.

In Section 2.7 we have met the sequence of Fibonacci numbers 1, 1, 2, 3, 5, 8, 13, . . . The terms of this sequence are defined by the recursive formula (or recurrence relation) $F(k + 1) = F(k - 1) + F(k)$. It is clearly an increasing sequence which tends to infinity. In the following example, we shall study a convergent sequence whose terms are defined by a recursive formula.

6.9.10. *Example.* Consider the sequence $\langle z_n \rangle$ whose terms are defined recursively by

$$z_1 = 4 \text{ and } z_{n+1} = \frac{1}{2}\left(z_n + \frac{4}{z_n}\right) \qquad (n = 1, 2, \ldots).$$

We observe that all terms of the sequence $\langle z_n \rangle$ are positive and by the recurrence relation above, they satisfy the following equation:

$$z_n^2 - 2z_{n+1}\, z_n + 4 = 0, \quad n = 1, 2, \ldots$$

In other words, such quadratic equations in the unknown z_n have positive solutions. Therefore, the discriminants of these equations are non-negative.

i.e. $\qquad (2z_{n+1})^2 - 4^2 \geqslant 0$

$\therefore \qquad z_{n+1} \geqslant 2, \quad$ for $n = 1, 2, \ldots$.

Hence the sequence $\langle z_n \rangle$ is bounded below by 2.

Next we wish to prove that $\langle z_n \rangle$ is a decreasing sequence. Now for all $n = 1, 2, \ldots$.

$$z_n - z_{n+1} = z_n - \frac{1}{2}\left(z_n + \frac{4}{z_n}\right)$$

$$= \frac{1}{2z_n}\left(z_n^2 - 4\right) \geqslant 0.$$

Therefore the sequence $\langle z_n \rangle$ and so also the sequence $\langle z_{n+1} \rangle$ converge to the infimum k of the set $S = \{z_n \mid n = 1, 2, \ldots\}$ of the sequence. Using the equation

$$z_n^2 - 2z_{n+1}\, z_n + 4 = 0$$

and the sum rule as well as the product rule of convergent sequences, we get

$$k^2 - 2k^2 + 4 = 0.$$

Therefore $k = 2$ or $k = -2$. But -2 cannot be the infimum of S since S is already known to be bounded below by 2. Therefore 2 is the infimum of S and hence $z_n \to 2$ as $n \to \infty$.

6.10. CAUCHY'S CONVERGENCE TEST

The two theorems of the last section provide a criterion of convergence for monotone sequences, which is without reference to a limit. In this section, we shall study a general criterion of convergence of all sequences which is expressed through the terms of the sequence alone.

Now, if a sequence $\langle a_n \rangle$ converges, then its terms will eventually be confined to every arbitrarily small neighbourhood of the limit. Or without reference to the limit, we may say that for sufficiently large indices m and n, the distance between a_m and a_n can be made as small as we like. More precisely, a necessary condition for the convergence of the sequence $\langle a_n \rangle$ is that *given any* $\epsilon > 0$ *there is a number* N *such that* $|a_n - a_m| < \epsilon$, *provided that* m *and* n *are both greater than* N.

Indeed if $\langle a_n \rangle \to a$ and $\epsilon > 0$ is a given positive number, then for $\epsilon/2 > 0$, we have a number N such that $|a_n - a| < \epsilon/2$ for all $n > N$. Now if $m, n > N$, then

$$\begin{aligned} |a_n - a_m| &= |(a_n - a) - (a_m - a)| \\ &\leqslant |a_n - a| + |a_m - a| \\ &< \epsilon, \end{aligned}$$

proving the above condition to be necessary for the convergence of $\langle a_n \rangle$ (Fig. 6.8).

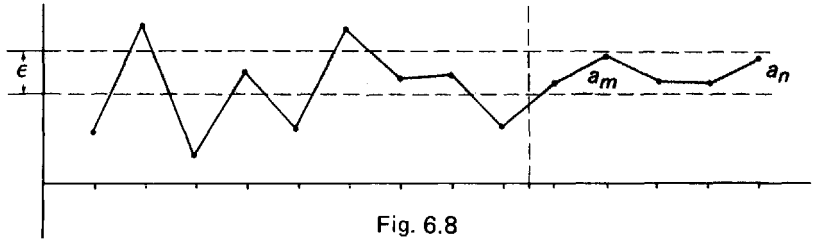


Fig. 6.8

The French mathematician Augustin-Louis Cauchy (1789–1857), who had contributed much to the foundation of modern mathematical analysis, proved that the above intrinsic condition is also sufficient for the convergence of the sequence $\langle a_n \rangle$. We formulate this important result in the theorem below.

6.10.1. Cauchy's Convergence Test. *Let $\langle a_n \rangle$ be a sequence. If for every $\epsilon > 0$, there is a number N such that*

$$|a_n - a_m| < \epsilon$$

whenever both indices m and n are greater than N, then the sequence $\langle a_n \rangle$ is convergent and possesses a limit.

Intuitively, the idea of Cauchy's test is quite simple if we consider the diagram above; however its proof requires some rather sophisticated techniques, which are beyond the scope of this book. Since we shall only use it in a few instances in the subsequent sections, we may simply accept its validity. In the appendix of this chapter, we shall, for the sake of completeness, give a proof of the test. Meanwhile, we shall use the test to obtain further examples of convergent sequences.

6.10.2. *Example.* Consider the sequence $\langle a_n \rangle$ whose terms are given by the following initial values and the recurrence relation

$$a_1 = 0, \quad a_2 = 1 \quad \text{and} \quad a_{n+1} = \frac{1}{2}(a_{n-1} + a_n) \quad \text{for } n = 2, 3, \ldots$$

In other words, every term of the sequence is the arithmetic mean of the preceding two terms. Therefore, the sequence $\langle a_n \rangle$ is not a monotone sequence. Though it is easily seen to be bounded by 0 and 1, the criteria for convergence of Theorems 6.9.6 and 6.9.7 are not applicable to the present sequence. Let us now examine the sequence by Cauchy's test. For the distance between two consecutive terms, we have

$$|a_2 - a_1| = 1, \; |a_3 - a_2| = \frac{1}{2}, \; |a_4 - a_3| = \frac{1}{4}, \ldots, \; |a_{n+1} - a_n| = \frac{1}{2^{n-1}}$$

If $m > n + 1$, then it is easy to see that the term a_m lies between a_n and a_{n+1}. With the distance between consecutive terms decreasing every step by half and with all following terms lying in between, we see that the sequence $\langle a_n \rangle$ satisfies the condition of Cauchy's test. Therefore $\langle a_n \rangle$ is

convergent. Moreover, we can verify by induction that

$$a_n = \frac{2}{3}\left(1 + \frac{(-1)^n}{2^{n-1}}\right).$$

Therefore $a_n \to 2/3$ as $n \to \infty$.

6.10.3. *Example.* Take the sequence $\langle x_n \rangle$ of Example 6.9.8 where

$$x_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n} \quad (n = 1, 2, \ldots)$$

Then for every n we have

$$|x_{2n} - x_n| = \left| \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} \right|$$
$$> n \cdot \frac{1}{2n} = \frac{1}{2}.$$

Therefore, for any $\epsilon < 1/2$, no number N can be found that can satisfy the condition of Cauchy's test. Therefore $\langle x_n \rangle$ is divergent.

6.10.4. *Example.* Consider the sequence $\langle y_n \rangle$ whose n-th term is given by an alternating sum

$$y_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{(-1)^{n-1}}{n}$$

Thus $y_1 = 1, y_2 = \frac{1}{2}, y_3 = \frac{5}{6}, y_4 = \frac{7}{12}, \ldots$

Again, $\langle y_n \rangle$ is not a monotone sequence. We shall show that it satisfies the condition of Cauchy's test.

For any two indices $m > n$, we have

$$|y_m - y_n| = \left| \frac{1}{n+1} - \frac{1}{n+2} + \frac{1}{n+3} - \frac{1}{n+4} + \ldots + (-1)^{m-n+1} \frac{1}{m} \right|$$

Therefore if $m - n$ is even, then

$$|y_m - y_n| = \left(\frac{1}{n+1} - \frac{1}{n+2}\right) + \ldots + \left(\frac{1}{m-1} - \frac{1}{m}\right)$$
$$= \frac{1}{n+1} - \left(\frac{1}{n-2} - \frac{1}{n-3}\right) - \ldots - \left(\frac{1}{m-2} - \frac{1}{m-1}\right) - \frac{1}{m}$$

and if $m - n$ is odd, then

$$|y_m - y_n| = \frac{1}{n+1} - \left(\frac{1}{n-2} - \frac{1}{n-3}\right) - \ldots - \left(\frac{1}{m-1} - \frac{1}{m}\right).$$

Thus in both cases we have

$$|y_m - y_n| < \frac{1}{n+1}.$$

Therefore $|y_m - y_n| < \epsilon$ as soon as both m and n are greater than $1/\epsilon$. Hence $\langle y_n \rangle$ is a convergent sequence.

6.11. EXERCISE

1. If $x_n = n/(n^2 + 1)$ for any positive integer n, prove that the sequence $\{x_n\}$ is increasing.

2. If $x_n = (2n - 7)/(3n + 2)$ for any positive integer n, show that the sequence $\{x_n\}$ is increasing.

3. (a) Suppose a sequence $\{x_n\}$ increases and is unbounded above. Prove that $x_n \to +\infty$ as $n \to \infty$.
   (b) If $\{x_n\}$ decreases and is unbounded below, prove that $x_n \to -\infty$ as $n \to \infty$.

4. Prove that a convergent sequence is always bounded.

5. If $\{a_n\}$ is an increasing and unbounded sequence of natural numbers, and if $\lim_{n\to\infty} b_n = l$, prove that $\lim_{n\to\infty} b_{a_n} = l$.

6. If a sequence $\{a_n\}$ of real numbers is strictly increasing, show that the sequence $\{b_n\}$ defined by
$$b_n = \frac{\sum_{i=1}^{n} a_i}{n}$$
is also strictly increasing.

7. If $a > 0$, $0 < r_0 < \sqrt{a_0}$ and for all $n \geqslant 0$,
$$r_{n+1} = \frac{r_n(3a + r_n)^2}{a + 3r_n{}^2},$$
prove that
$$a - r_{n+1}{}^2 = \frac{(a - r_n{}^2)^3}{a + 3r_n{}^2}$$
and
$$r_{n+1} - r_n = \frac{2(a - r_n{}^2)r_n}{a + 3r_n{}^2}.$$
Deduce that r_n is monotonic increasing and bounded, and that $\lim_{n\to\infty} r_n = \sqrt{a}$.

8. (a) By using the inequality A.M. $\geqslant$ G.M. or otherwise, prove that
$$(1+\frac{1}{n-1})^{n-1} \leqslant (1+\frac{1}{n})^n$$
for any positive integer $n \geqslant 2$.
(b) Show that the sequence $\{(1+1/n)^n\}$ converges to a limit $\leqslant 3$.

*9. If $0 < u_1 < 3$ and $u_{n+1} = 12/(1+u_n)$, show that the sequences $\{u_{2n+1}\}$ and $\{u_{2n}\}$ are respectively monotonic increasing and decreasing. Show also that the sequence $\{u_n\}$ converges to the limit 3.

6.12. SERIES

Given a sequence $\langle a_n \rangle$, we can always obtain a new sequence
$$a_1, \quad a_1+a_2, \quad a_1+a_2+a_2, \quad a_1+a_2+a_3+a_4, \ldots$$
by adding up successive terms. If we denote the N-th term of the new sequence by s_N, $N = 1, 2, \ldots$, then we have
$$s_N = a_1 + a_2 + \ldots + a_N.$$
The new sequence $\langle s_N \rangle$ of partial sums is called an *infinite series* or simply a *series*. We shall call the number a_n the n-th *term* and the sum s_N a *partial sum* of the series. The series is also written in any one of the following notations:
$$a_1 + a_2 + a_3 + \ldots.$$
$$a_1 + a_2 + \ldots + a_n + \ldots$$
$$\sum_{n=1}^{\infty} a_n$$

We remark immediately that the above three notations must be interpreted as complexes of symbols to denote the sequence $\langle s_N \rangle$. They should not be taken to denote 'infinite sums' because addition is defined for two summands and can be extended only inductively for a finite number of summands. Thus
$$1+\frac{1}{2}+\frac{1}{3}+\ldots$$
$$1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+\ldots$$
$$\sum_{n=1}^{\infty}\frac{1}{n}$$
all represent the same sequence of the partial sums $\sum_{n=1}^{N}\frac{1}{n}$.

We say that the series $\Sigma_{n=1}^{\infty} a_n$ *converges to the sum* s if the sequence $\langle s_N \rangle$ of partial sums converges to the limit s. In this case we may also say that s is *the sum of the series* $\Sigma_{n=1}^{\infty} a_n$. Notice that the sum of a convergent series is a number which is the limit of the convergent sequence of partial sums. Therefore, it should not be confused with the series $\Sigma_{n=1}^{\infty} a_n$ itself, and it is not obtained by ordinary addition of the terms a_n. However it is customary and convenient to write $s = \Sigma_{n=1}^{\infty} a_n$ if the series $\Sigma_{n=1}^{\infty} a_n$ converges to sum s, abusing the proper meaning of the equality sign.

We have seen earlier in Examples 6.9.8, 6.9.9 and 6.10.4 that the series $\Sigma_{n=1}^{\infty} 1/n$ is divergent while the series $\Sigma_{n=1}^{\infty} 1/n!$ and the series $\Sigma_{n=1}^{\infty} (-1)^{n-1}/n$ are both convergent. In particular, the base e of natural logarithm is the sum of the convergent series $\Sigma_{n=1}^{\infty} 1/n!$

6.12.1. Theorem. *If a series*

$$a_1 + a_2 + \ldots + a_n + \ldots$$

converges to a sum s, *then the sequence* $\langle a_n \rangle$ *of terms of the series is a null sequence.*

Proof. If the series $a_1 + a_2 + a_3 + \ldots$ converges to s, then the sequence of partial sums $\langle s_N \rangle$ converges to s by definition. Now $a_N = s_{N+1} - s_N$. By the sum rule, we have $\lim a_N = \lim s_{N+1} - \lim s_N = s - s = 0$. Therefore $\langle a_n \rangle$ is a null sequence. ∎

Obviously the convergence of a series does not necessarily follow from the fact that its terms a_n tend to zero as n increases. A typical counter example is provided by the so called harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots$$

which diverges while its term $1/n$ converges to 0. On the other hand, it follows from Theorem 6.12.1 that if the terms b_n of a series $\Sigma_{n=1}^{\infty} b_n$ do not tend to 0 as $n \to \infty$, then the series is divergent.

6.13. GEOMETRIC SERIES AND HARMONIC SERIES

A *geometric series* is obtained by adding together successive terms of a geometric progression

$$1, x, x^2, x^3, \ldots.$$

and is denoted by

$$1 + x + x^2 + \ldots + x^n + \ldots \quad \text{or} \quad \sum_{n=0}^{\infty} x^n$$

Notice here that for convenience, we begin the series with the index $n = 0$. The partial sum s_N is given by

$$s_N = 1 + x + x^2 + \ldots + x^N.$$

If $x = 1$, then $s_N = N + 1$. Clearly $s_N \to \infty$ as $N \to \infty$. Thus the series $\Sigma_{n=0}^{\infty} x^n$ is divergent for $x = 1$. If $x \neq 1$, then

$$s_N = 1 + x + x^2 + \ldots + x^N = (1 - x^{N+1})/(1 - x).$$

Thus the convergence of the geometric series depends entirely on the value of x. For $|x| < 1$, we know that $x^{N+1} \longrightarrow 0$ as $n \longrightarrow \infty$. For $|x| > 1$, the sequence $\langle 1 - x^{N+1} \rangle$ clearly diverges. Therefore we have proved the following theorem on geometric series.

6.13.1. Theorem. *The geometric series*

$$1 + x + x^2 + \ldots + x^n + \ldots$$

converges to the sum $1/(1 - x)$ *if* $|x| < 1$, *and diverges if* $|x| \geqslant 1$.

The *harmonic series*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

is the sequence of partial sums of the null sequence $\langle 1/n \rangle$. We have seen in Example 6.9.8 that the harmonic series is a divergent series. A generalization of the harmonic series is the series

$$\sum_{n=1}^{\infty} \frac{1}{n^k} = 1 + \frac{1}{2^k} + \frac{1}{3^k} + \ldots$$

where k is a fixed real number. For such series we can prove the following theorem.

6.13.2. Theorem. *The general harmonic series*

$$\sum_{n=1}^{\infty} \frac{1}{n^k} = 1 + \frac{1}{2^k} + \frac{1}{3^k} + \ldots$$

converges if $k > 1$ *and diverges if* $k \leqslant 1$.

Proof. (a) $k = 1$. Here we have the ordinary harmonic series, and we have seen that it diverges.

(b) $k < 1$. For the partial sums we have

$$s_N = 1 + \frac{1}{2^k} + \ldots + \frac{1}{N^k} > 1 + \frac{1}{2} + \ldots + \frac{1}{N}$$

As N increases, the right-hand side tends to infinity and so $s_N \to \infty$ also.

(c) $k > 1$. Since all terms a_n of the series are positive, the sequence $\langle s_N \rangle$ of partial sums is a strictly increasing sequence. We shall use the familiar device of grouping terms in blocks of 2, 4, 8, . . . summands to show that the sequence $\langle s_N \rangle$ is bounded above. We have

$$\frac{1}{2^k} + \frac{1}{3^k} < \frac{2}{2^k} = \frac{1}{2^{k-1}}$$

$$\frac{1}{4^k} + \frac{1}{5^k} + \frac{1}{6^k} + \frac{1}{7^k} < \frac{4}{4^k} = \frac{1}{4^{k-1}}$$

$$\frac{1}{8^k} + \frac{1}{9^k} + \ldots + \frac{1}{15^k} < \frac{8}{8^k} = \frac{1}{8^{k-1}}$$

. . . . . .

Therefore for $M = 2^N - 1$, we have

$$s_M = 1 + \frac{1}{2^k} + \ldots + \frac{1}{M^k} < 1 + \frac{1}{2^{k-1}} + \frac{1}{4^{k-1}} + \ldots + \frac{1}{(2^{N-1})^{k-1}}.$$

The right-hand side is the sum of a finite geometric progression which is equal to

$$\frac{(1 - \frac{1}{2^{N(k-1)}})}{(1 - \frac{1}{2^{k-1}})}.$$

Therefore

$$s_M < 1/(1 - 2^{1-k})$$

proving that the increasing sequence $\langle s_N \rangle$ is bounded above. Hence by Theorem 6.9.6, $\langle s_N \rangle$ converges. ∎

As with most infinite series, we do not have a simple expression for the sum of the general harmonic series.

6.14. SOME USEFUL RULES

Some of the rules for convergent sequences can be easily written into rules for convergent series. We list below five useful rules which can be easily proved.

1. *The convergence and divergence of a series is unaffected if a finite number of terms are inserted, suppressed or altered.*
2. *If* $\Sigma_{n=1}^{\infty} a_n$ *converges to the sum s and* $\Sigma_{n=1}^{\infty} b_n$ *converges to the sum t, then the series*

$$\sum_{n=1}^{\infty} (\lambda a_n + \mu b_n) = (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) + \ldots$$

converges to the sum $\lambda s + \mu t$ *for any real numbers* λ *and* μ.
3. *If* $a_n \geqslant 0$ *for* $n = 1, 2, \ldots$, *then the series* $\Sigma_{n=1}^{\infty} a_n$ *either converges or diverges to infinity. A necessary and sufficient condition for convergence is that there is a fixed real number K such that the partial sums* $a_1 + \ldots + a_N < K$ *for* $N = 1, 2, 3, \ldots$; *and in this case the sum of the series is less than or equal to K.*
4. *Let* $\Sigma_{n=1}^{\infty} a_n$ *and* $\Sigma_{n=1}^{\infty} b_n$ *be two series such that for all* $n = 1, 2, \ldots$, $0 \leqslant a_n \leqslant \mu b_n$ *for some constant* μ. *If* $\Sigma_{n=1}^{\infty} b_n$ *converges to the sum t then* $\Sigma_{n=1}^{\infty} a_n$ *converges to a sum which is less than or equal to* μt.
5. *Let* $\Sigma_{n=1}^{\infty} x_n$, $\Sigma_{n=1}^{\infty} a_n$ *and* $\Sigma_{n=1}^{\infty} y_n$ *be three series. If* $x_n \leqslant a_n \leqslant y_n$ *for all* $n = 1, 2, \ldots$ *and if* $\Sigma_{n=1}^{\infty} x_n$ *and* $\Sigma_{n=1}^{\infty} y_n$ *converge to the same sum s, then* $\Sigma_{n=1}^{\infty} a_n$ *converges to the sum s.*

6.15. TEST OF CONVERGENCE

The rules of Section 6.14 provide us with a variety of ways to test the convergence of a series by means of comparison with known convergent series. An intrinsic criterion for convergence of series is obtained by an easy adoptation of Cauchy's convergence test.

6.15.1. Theorem. *A necessary and sufficient condition for the convergence of a series* $\Sigma_{n=1}^{\infty} a_n$ *is that given any* $\epsilon > 0$, *there is a number N such that*

$$|a_{n+1} + a_{n+2} + \ldots + a_m| < \epsilon$$

provided that $N < n < m$.

As an application of the above, we shall derive from it a handy criterion for the convergence of alternating series. These are series of the form

$$\sum_{n=1}^{\infty} (-1)^{n-1} a_n = a_1 - a_2 + a_3 - \ldots .$$

6.15.2. *Example.* If $\langle a_n \rangle$ is a decreasing sequence of positive numbers with $\lim a_n = 0$, then the alternating series

$$a_1 - a_2 + a_3 - a_4 + \ldots + (-1)^{n-1} a_n + \ldots .$$

is convergent.

Proof. For a pair of indices $n < m$, we consider

$$T = a_{n+1} - a_{n+2} + a_{n+3} - \ldots + (-1)^{m-n+1} a_m .$$

Depending on whether $m - n$ is even or odd, we get

$$T = (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \ldots + (a_{m-1} - a_m)$$

or $$T = (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \ldots + (a_{m-2} - a_{m-1}) + a_m .$$

Since $a_n \searrow$ we get T non-negative in both cases. On the other hand, also according to whether $m - n$ is even or odd, we have

$$T = a_{n+1} - (a_{n+2} - a_{n+3}) - \ldots - (a_{m-2} - a_{m-1}) - a_m$$

or $$T = a_{n+1} - (a_{n+2} - a_{n+3}) - \ldots - (a_{m-1} - a_m) .$$

Therefore, in both cases, $T \leqslant a_{n+1}$. It now follows from these two interim results that for all $n < m$

$$|(-1)^n a_{n+1} + (-1)^{n+1} a_{n+2} + \ldots + (-1)^{m-1} a_m| = T \leqslant a_{n+1} .$$

Since $a_{n+1} \longrightarrow 0$ as $n \longrightarrow \infty$, the expression of the left-hand side can be made arbitrarily small for sufficiently large $n < m$. Therefore by Theorem 6.15.1, the alternating series is convergent. ∎

6.16. EXERCISE

1. Evaluate $\lim\limits_{n\to\infty} \Sigma_{i=n}^{2n} \frac{1}{i^2}$.

2. Show that the series $\Sigma_{n=1}^{\infty} \frac{n}{(n+1)!}$ converges to 1.

3. Show that the series $\Sigma_{n=1}^{\infty} (-1)^{n+1} (2n - 1)$ is divergent.

4. Show that the series $\Sigma_{n=1}^{\infty} \frac{1}{(3n-1)(3n+2)}$ converges to $\frac{1}{6}$.

5. Sum to n terms the series

$$\frac{1}{1 \times 3 \times 5} + \frac{1}{3 \times 5 \times 7} + \ldots + \frac{1}{(2n+1)(2n+3)(2n+5)} + \ldots$$

and deduce the sum to infinity.

6. Sum to n terms the series whose n-th term is $1/[n(n+1)(n+3)]$ and deduce the sum to infinity.

7. If $a_n > 0$ and $b_n > 0$ $(n = 1, 2, 3, \ldots)$, and $\lim_{n\to\infty} (a_n/b_n) = c \neq 0$, prove that $\Sigma_{n=1}^{\infty} a_n$ converges if and only if $\Sigma_{n=1}^{\infty} b_n$ converges.

8. Suppose a sequence $\{a_n\}$ is decreasing and $\lim_{n\to\infty} a_n = 0$. If $\Sigma_{n=1}^{\infty} a_n$ converges, prove that $\Sigma_{n=1}^{\infty} 2^n a_{2^n}$ also converges. (Cauchy Condensation Theorem).

9. If $a_n \geqslant 0$ $(n = 1, 2, 3, \ldots)$ and $\lim_{n\to\infty} \sqrt[n]{a_n} = r$, prove that $\Sigma_{n=1}^{\infty} a_n$ converges if $r < 1$ and diverges if $r > 1$. (This result is known as the 'root test'.)

10. Suppose $\{b_n\}$ is a decreasing sequence, with $b_n \geqslant 0$ for every positive integer n. If $m \leqslant \Sigma_{i=1}^{n} a_i \leqslant M$ for any positive integer n, prove that
$$b_1 m \leqslant \Sigma_{i=1}^{n} a_i b_i \leqslant b_1 M. \quad \text{(Abel's Lemma)}$$
(Hint: $a_i = \Sigma_{j=1}^{i} a_j - \Sigma_{j=1}^{i-1} a_j$)

11. (a) Express $\sin 3\theta$ in terms of $\sin\theta$.
    (b) Find the sum to n terms and the sum to infinity of the series whose n-th term is $3^{n-1} \sin \dfrac{\theta}{3^n}$.

*12. (a) Let $\{a_n\}$ be a sequence of integers with $0 \leqslant a_n \leqslant 9$. Prove that $\Sigma_{n=1}^{\infty} a_n 10^{-n}$ exists and lies between 0 and 1. (This, of course, is the number which we usually denote by $0.a_1a_2a_3\ldots$.)
    (b) Suppose $0 \leqslant x \leqslant 1$. Prove that there is a sequence of integers $\{a_n\}$ with $0 \leqslant a_n \leqslant 9$ and
$$\Sigma_{n=1}^{\infty} a_n 10^{-n} = x.$$
(Hint: For example, $a_1 = [10x]$, where $[y]$ denotes the greatest integer which is not greater than y.)
    (c) Show that if $\{a_n\}$ is repeating, i.e. is of the form $a_1, a_2, \ldots, a_k, a_1, a_2, \ldots, a_k, a_1, a_2, \ldots, a_k, \ldots$, then $\Sigma_{n=1}^{\infty} a_n 10^{-n}$ is a rational number and find it.
    (d) If $x = \Sigma_{n=1}^{\infty} a_n 10^{-n}$ is a rational number, prove that $\{a_n\}$ is eventually repeating.

13. If $a_n > 0$ for any positive integer n, and the sequence $\{a_n\}$ is strictly decreasing and $\Sigma_{n=1}^{\infty} a_n$ is convergent, prove that $\lim_{n\to\infty} na_n = 0$.

6.17. APPENDIX

In this appendix, we shall present a proof of Cauchy's test for convergence (see Theorem 6.10.1). The proof is based on the very important postulate of continuity (see Theorem 5.7.2). If the reader finds the presentation natural and comprehensible, he can congratulate himself. If he finds it too difficult to follow, he should not be unduly worried, but should leave it and read the next chapter. The arguments that we use here are rather above the level of an A-level course.

A sequence $\langle a_n \rangle$ is called a *Cauchy sequence* or a *fundamental sequence* if the following condition is satisfied:

Given any $\epsilon > 0$, there is a number N such that $|a_n - a_m| < \epsilon$ for all indices m and n greater than N.

6.17.1. Theorem. *A sequence $\langle a_n \rangle$ converges to a limit if and only if it is a Cauchy sequence.*

Proof. We observe that given a Cauchy sequence, we can find for every $\epsilon > 0$ an index p such that

$$|a_n - a_p| < \epsilon \text{ for all } n > p.$$

In other words, the interval $(a_p - \epsilon, a_p + \epsilon)$ will contain the terms $a_{p+1}, a_{p+2}, \ldots$ of the sequence $\langle a_n \rangle$.

Suppose that $\langle a_n \rangle$ is a Cauchy sequence. Let $\epsilon_1 = 1/2$, $\epsilon_2 = 1/4, \ldots$, $\epsilon_n = 1/2^n, \ldots$. Then it follows from the above observation that for ϵ_1, we can find an index p_1 such that

$$|a_n - a_{p_1}| < \frac{1}{2} \text{ for all } n > p_1.$$

Similarly for $\epsilon_2 = 1/4$, we can find among the indices greater than p_1 an index p_2 such that

$$|a_n - a_{p_2}| < \frac{1}{2^2} \text{ for all } n > p_2 > p_1.$$

Inductively for every $k = 1, 2, \ldots$, we obtain an index p_k such that

$$p_1 < p_2 < p_3 \ldots < p_k < \ldots.$$

and

$$|a_n - a_{p_k}| < \frac{1}{2^k} \text{ for all } n > p_k.$$

If we denote by I_k the interval $(a_{p_k} - \epsilon_k, a_{p_k} + \epsilon_k)$, then these intervals have the properties

(i) the length of I_k is $1/2^{k-1}$;

(ii) the midpoint of I_k is the term a_{p_k} of the sequence $\langle a_n \rangle$;

(iii) $a_n \in I_k$ for all $n \geqslant p_k$, i.e. a 'tail' of the sequence $\langle a_n \rangle$ falls into I_k.

Though I_k is only half the length of the preceding I_{k-1}, the relative positions of their midpoints may not be such that $I_k \subset I_{k-1}$. However, I_k and I_{k-1} do overlap in any case, since the midpoint of I_k lies in I_{k-1}, as well as in all the preceding intervals $I_1, I_2, \ldots, I_{k-2}$. This allows us to cut away certain portion from each I_k, so that the 'tailored' intervals J_k will be nested in one another and still contain the same terms of the sequence $\langle a_n \rangle$ as the original intervals I_k. More precisely, we define

$$J_1 = I_1$$
$$J_2 = J_1 \cap I_2$$
$$\ldots\ldots\ldots\ldots$$
$$J_k = J_{k-1} \cap I_k$$
$$\ldots\ldots\ldots\ldots$$

Now the 'tailored' intervals J_k have the following properties:

(iv) $J_1 \supset J_2 \supset \ldots \supset J_k \supset \ldots$ i.e. they are nested in one another;

(v) the length of J_k is less than or equal to $1/2^{k-1}$ but greater than 0;

(vi) $a_{p_k} \in J_k$;

(vii) $a_n \in J_k$, for all $n \geqslant p_k$.

If we denote by x_k and y_k the endpoints of J_k, i.e. $J_k = (x_k, y_k)$, then

$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_k \leqslant \ldots \leqslant y_k \leqslant \ldots \leqslant y_2 \leqslant y_1$$

and

$$|x_k - y_k| \leqslant \frac{1}{2^{k-1}}.$$

Therefore the sequence $\langle x_n \rangle$ is an increasing sequence which is bounded above, and $\langle y_n \rangle$ is a decreasing sequence bounded below. They both converge. On the other hand, $|x_k - y_k| \to 0$ as $n \to \infty$. The supremum of the set $\{x_k \mid k = 1, 2, \ldots\}$ must be identical with the infimum of the set $\{y_k \mid k = 1, 2, \ldots\}$; therefore the sequences $\langle x_n \rangle$ and $\langle y_n \rangle$ converge to the same limit a. Thus

$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_k \leqslant \ldots \leqslant a \leqslant \ldots \leqslant y_k \leqslant \ldots \leqslant y_2 \leqslant y_1.$$

We now claim that $\langle a_n \rangle$ converges to the limit a. Suppose $\epsilon > 0$. Then

we can find an index k such that $\epsilon > 1/2^k$. It follows from (iii) that for all $n > p_k$

$$|a_n - a_{p_k}| < \frac{1}{2^k}.$$

On the other hand, both a_{p_k} and a belong to the interval J_k; therefore

$$|a - a_{p_k}| < \frac{1}{2^{k-1}}.$$

Combining these two inequalities, we get for all $n > p_k$,

$$\begin{aligned} |a_n - a| &= |(a_n - a_{p_k}) - (a - a_{p_k})| \\ &\leqslant |a_n - a_{p_k}| + |a - a_{p_k}| \\ &< \frac{1}{2^{k-1}} + \frac{1}{2^{k-1}} < \epsilon \end{aligned}$$

Therefore $\lim a_n = a$ proving that the Cauchy sequence $\langle a_n \rangle$ is convergent. We have seen in Section 6.10 that every convergent sequence is a Cauchy sequence. The proof of the theorem is now complete. ■

7. Complex Numbers

Instead of starting with a formal definition of complex number and then proceeding to study the properties of the new number system, we shall begin with a brief review of our old number systems $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ in order to discover a certain inadequacy in each concerning the provision of solutions of equations. This inadequacy will be partially overcome by successive extensions. Ultimately the new number system $\mathbb{C}$ of complex members will be seen as the final result of the effort in removing this inadequacy.

7.1. EQUATIONS AND NUMBER SYSTEMS

An equation can be regarded as a mathematical expression of a certain condition or specification on an unknown number. Take, for instance, an example from primary school mathematics. John has saved a sum of money. If his father gives him 45 dollars, he will have altogether a total of 200 dollars. How much has John saved? The condition that 45 added to the unknown number x of dollars that John has saved will give 200 can be written as a simple equation

$$x + 45 = 200.$$

We may say that this equation is *formulated in the number system* $\mathbb{N}$ in the sense that all its numbers belong to $\mathbb{N}$ and all its symbols have meaning in the system $\mathbb{N}$. The equation admits a solution in the system $\mathbb{N}$ because we can find a number 155 in $\mathbb{N}$ such that $155 + 45 = 200$. Similarly

$$x + 50 = 214, \quad x + 7 = 83, \quad 5y = 25$$

are equations which are formulated in $\mathbb{N}$ and admit solutions in $\mathbb{N}$.

However, it is not always true that an equation which is formulated in $\mathbb{N}$ would admit a solution in $\mathbb{N}$. Take, for instance, the equation

$$x + 45 = 5.$$

This can well be the mathematical expression of a real life situation. For example, after I put 45 dollars into my account, the balance will show 5 dollars and I wish to know the balance before that. The equation is formulated in $\mathbb{N}$ but admits no solution in $\mathbb{N}$, because there is no natural

number that will give 5 when added to 45. The equations

$$2x + 3 = 2, \quad x^2 + 5x + 6 = 0, \quad x^2 = 2, \quad x^2 + 1 = 0$$

are all formulated in $\mathbb{N}$ but admit no solution in $\mathbb{N}$. Undoubtedly this shows that $\mathbb{N}$ is too small to admit solutions of all equations that are formulated in $\mathbb{N}$.

This inadequancy of $\mathbb{N}$ can be overcome, at least partially, by extending $\mathbb{N}$ to the next larger system $\mathbb{Z}$ of intergers. By saying that the system $\mathbb{Z}$ is an extension of the system $\mathbb{N}$, we mean not only that the set $\mathbb{Z}$ contains the set $\mathbb{N}$ as a proper subset, but also that addition and multiplication are extended as well. Under the extension. equations which are formulated in $\mathbb{N}$ remain formulated in $\mathbb{Z}$. Now some equations that are formulated in $\mathbb{N}$ but admit no solution in it will admit solutions in $\mathbb{Z}$. For example, $x + 45 = 5$ admits the solution -40 in $\mathbb{Z}$ and the quadratic equation $x^2 + 5x + 6 = 0$ admits two solutions -2 and -3 in $\mathbb{Z}$.

However, the system $\mathbb{Z}$ is still not large enough to admit solutions of the equation $2x + 3 = 2$ which is formulated in $\mathbb{Z}$. Similar to what is done before, this inadequancy is partially overcome when $\mathbb{Z}$ is extended to the number system $\mathbb{Q}$ of rational numbers. In $\mathbb{Q}$, $2x + 3 = 2$ admits the solution $-1/2$. However $x^2 = 2$ and $x^2 + 1 = 0$ still have no solutions in $\mathbb{Q}$. This means that we have to extend $\mathbb{Q}$ one step further to the number system $\mathbb{R}$ of real numbers. Then $x^2 = 2$ has solutions $\pm\sqrt{2}$ in $\mathbb{R}$ while $x^2 + 1 = 0$ which is formulated in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ as well as in $\mathbb{R}$ still admits no solutions in $\mathbb{R}$ since the squares of all real numbers are non-negative.

The successive extensions leading finally to the system $\mathbb{R}$ fail to remove the inadequancy entirely. There are still many equations, which are expressed in terms of real numbers and their addition and multiplication, but admit no solution in $\mathbb{R}$. This means that $\mathbb{R}$ has to be extended to some larger number system $\mathbb{C}$.

The set $\mathbb{C}$ should contain $\mathbb{R}$ as a proper subset; moreover, it must at least contain a solution of the equation $x^2 + 1 = 0$. Furthermore, addition and multiplication of numbers of the system should also extend those of real numbers, in the sense that the sum and the product of any two real numbers should be the same sum and the same product when regarded as numbers of the new system $\mathbb{C}$.

7.2. ONE-DIMENSIONAL NUMBER SYSTEM

As a preparation for our problem of extending the number system $\mathbb{R}$ to

a larger system $\mathbb{C}$, so that at least $x^2 + 1 = 0$ has a solution in $\mathbb{C}$, we first interpret the set $\mathbb{R}$ as a one-dimensional space and the algebraic operations of addition and multiplication as geometric operations of motions.

Following the method of Sections 5.1 and 5.7, we use a horizontal straight line to represent the set $\mathbb{R}$. After designating a point O as the number 0 and a segment as the unit length, we take the point A to the right of O at a distance of a unit lengths from O as the point representing the positive real number a. Similarly a point B to the left of O at a distance of b unit lengths from O represents the negative real number $-b$ (Fig. 7.1).

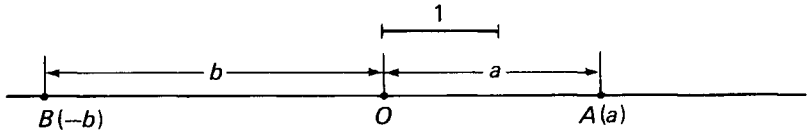


Fig. 7.1

In this way, all real numbers (i.e. including all rational numbers as well as irrational numbers) are points of a one-dimensional space — the number line. In this sense, we may say that the real numbers constitute a *one-dimensional number system.*

Suppose that two real numbers a and b are represented by the points A and B on the number line. If we move the point A a distance of $|b|$ (i.e. the length $|OB|$) in the direction from O to B (i.e. that of the directed segment OB), we shall arrive at the point C on the number line, which represents the sum $c = a + b$ (Fig. 7.2).

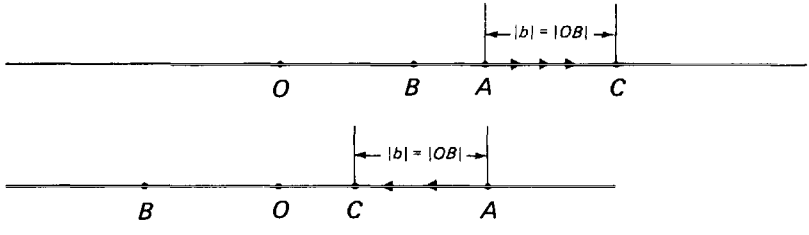


Fig. 7.2

Thus addition may be interpreted geometrically as parallel translation.

To locate the point D representing the product $d = ab$, we proceed in two steps. The first step is to get the point H, representing $h = |a|\,b$ by similarity (i.e. enlarging OB to OH by a factor equal to the length $|OA|$) as illustrated by Fig. 7.3 below.

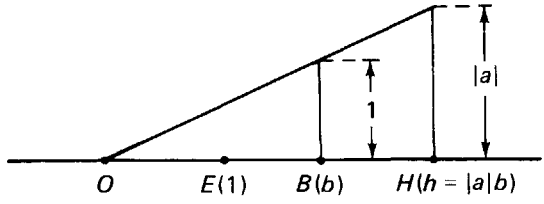


Fig. 7.3

To obtain the desired point D from the point H, we have to distinguish two cases: (i) $a > 0$ and (ii) $a < 0$.

For (i), $d = 1 \times h$. Therefore D is identical to H.

For (ii) $d = -1 \times h$. Therefore D is diametrically opposite to H with respect to O.

The appropriate geometric construction of D is therefore to rotate H about O by 0° or 360° in case (i) and to rotate H about O by 180° in case (ii), the angle of rotation being $\angle EOA$ in both cases, E being the point representing the number 1 (Fig. 7.4).

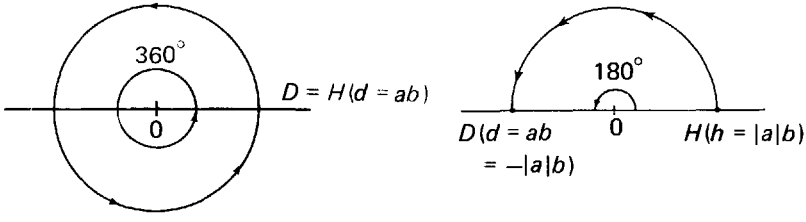


Fig. 7.4

Thus the multiplication of b by a may be interpreted geometrically as a similarity of OB by a factor $|OA|$ followed by a rotation about O through the angle $\angle EOA$. In particular, multiplication by -1 is a rotation about O by an angle $\pi = \angle EOF$, F being the point representing the number -1.

7.3. TWO-DIMENSIONAL NUMBER SYSTEM

We have seen in Section 7.1 that the equation

$$x^2 + 1 = 0$$

is formulated in the system $\mathbb{R}$ but fails to admit a solution in $\mathbb{R}$. Our final goal is to extend the system $\mathbb{R}$ (which includes the set $\mathbb{R}$ together with addition and multiplication) to some larger system $\mathbb{C}$ that has a solution of this equation which we tentatively denote by i. Some rough idea about a possible geometric picture of the extended system may help us reach the goal.

As the extended set $\mathbb{C}$ must contain $\mathbb{R}$ as a subset and the new number i as an element, we may extend the number line to a plane that contains the number line as the 2-dimensional space to represent the set $\mathbb{C}$. Under the additional but reasonable assumption that addition and multiplication of real numbers retain the geometric interpretations as described in the last section, we now try to find an appropriate position for the point P to represent the new number i that can give expression to the characteristic property of the number i.

As a solution of the equation

$$x^2 + 1 = 0$$

the number i is characterized by

$$i^2 = -1.$$

The distance to O from the point A representing a real number a is given by

$$|a| = \sqrt{a^2} = \sqrt{|a^2|}.$$

By analogy, we expect that the distance from P to O would be

$$\sqrt{|i^2|} = 1.$$

Therefore, the desired point P should be on the unit circle of radius 1, centred at O. If we denote by E the point on the real number line representing 1 and by θ the angle $\angle EOP$, then the angle θ determines the position of P (Fig. 7.5).

We now proceed to find θ. Consider now the two equations below:

$$-1 = -1 \times 1$$
$$-1 = i \times (i \times 1).$$

The geometric interpretation of the first equation is that by rotating the point E about O by 180°, we get the point F, representing the number -1 (Fig. 7.6). The right-hand side of the second equation consists of two multiplications,

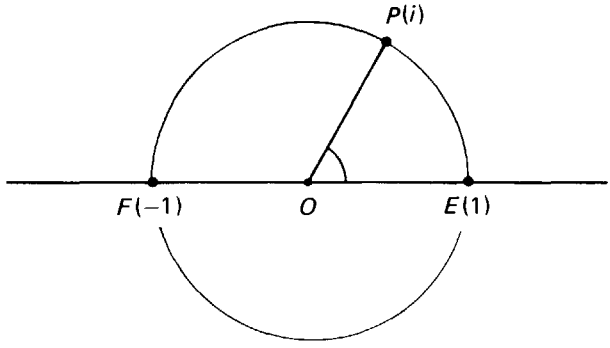


Fig. 7.5

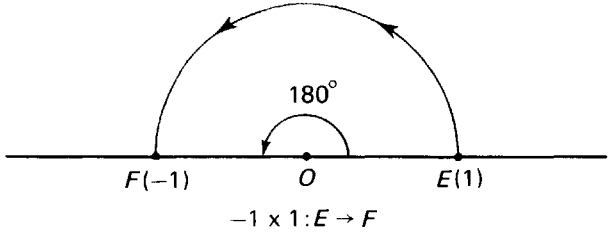


Fig. 7.6

i x 1 and then i x i. Adhering to the same interpretation, multiplication of 1 by i is an enlargment of OE by a factor $|OP| = 1$, followed by a rotation about O of θ (Fig. 7.7).

Now the enlargement of OE results in OE itself since $|OP| = 1$ and the rotation brings E to P. This is confirmed by $i = i$ x 1. Similarly the second multiplication rotates the point P about O by an angle $\theta = \angle EOP$. The result of this is the point R on the unit circle with $\angle EOR = 2\theta$ (Fig. 7.7).

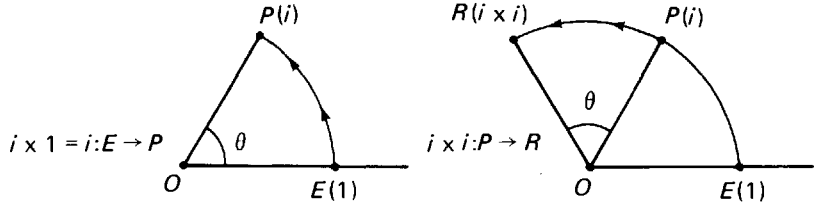


Fig. 7.7

On the other hand, R must be identical to F since $-1 \times 1 = i \times (i \times 1)$. Therefore $2\theta = 180^\circ$, and hence $\theta = 90^\circ$.

The point P representing the new number i is located in the diagram below (Fig. 7.8).

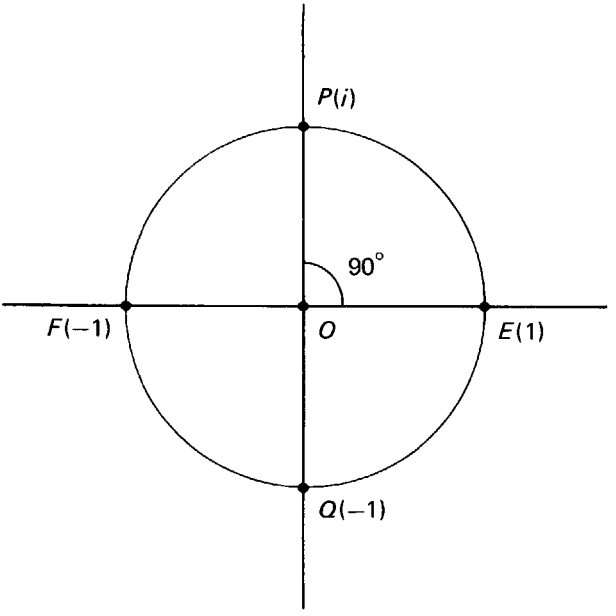


Fig. 7.8

The point Q diametrically opposite to P with respect to O can be obtained by rotating P about O by 180°. Q must therefore be the point representing the product $-i = -1 \times i$. Using the points on the real number line and the two new points P and Q, we can reach out at all other points on the plane by parallel translations and similarities. Re-interpreting points as numbers of the extended system $\mathbb{C}$, we see that the entire plane is a geometric representation of the two-dimensional number system $\mathbb{C}$.

Obviously our task is now to formulate this tentative geometric idea in precise algabraic terms. For this purpose we borrow the basic principle of analytic geometry to assign a pair of coordinates to each point on the plane.

7.4. COMPLEX NUMBERS

We define formally a *complex number* to be an ordered pair (a, b) of real numbers. The set of all complex numbers is denoted by the bold-faced capital letter $\mathbb{C}$. With the aid of a pair of coordinate axes we can represent every complex number $x = (a, b)$ by a point on the plane (Fig. 7.9).

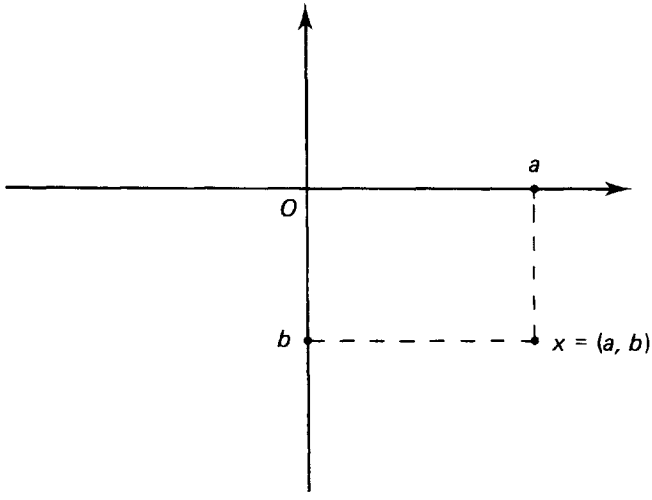


Fig. 7.9

Just as we sometimes refer to the number line as the *real line* or the *real axis* $\mathbb{R}$, we may call the plane of complex numbers the *complex plane* $\mathbb{C}$ or the *Gaussian plane* after the German mathematician Carl Friedrich Gauss (1777 – 1855).

It follows from the definition that two complex numbers $x = (a, b)$ and $y = (c, d)$ are equal if and only if they have identifical coordinates, i.e.

$x = y$ if and only if $a = c$ and $b = d$

$x \neq y$ if and only if $a \neq c$ or $b \neq d$.

The complex number $(0, 0)$ is called the zero complex number and is denoted by 0. Similarly if $x = (a, b)$ we denote the complex number $(-a, -b)$ by $-x$.

Addition of complex numbers. In order to make $\mathbb{C}$ into a fully fledged

number system, we have to define addition and multiplication. Following the preparation of Sections 7.2 and 7.3, we expect an addition of complex numbers to be interpreted as a parallel translation. In algebraic terms, parallel translation corresponds to coordinate-wise addition. Thus we define the sum $x + y$ of two complex numbers $x = (a, b)$ and $y = (c, d)$ to be the complex number

$$x + y = (a + c, b + d).$$

Fig. 7.10 confirms that the point C representing $x + y$ can be obtained by translating the directed segment OB along the directed segment OA where A and B represent x and y respectively.

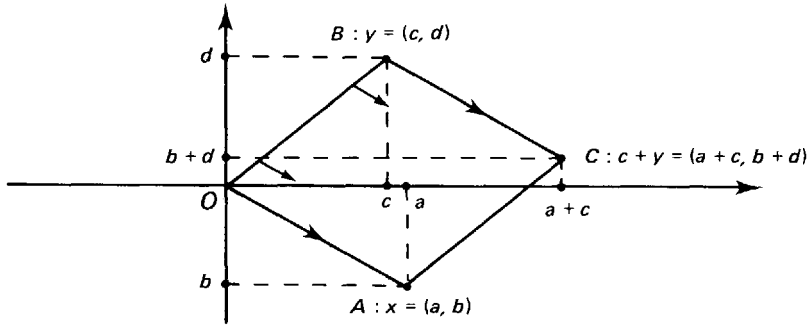


Fig. 7.10

Notice that $OACB$ is a parallelogram on the plane.

Obviously the *commulative law* of addition

$$x + y = y + x$$

and the *associative law* of addition

$$(x + y) + z = x + (y + z)$$

hold. Moreover, the *zero* complex number $0 = (0, 0)$ has the property that

$$0 + x = x \text{ for all } x \in \mathbb{C}.$$

Subtraction is defined as the inverse operation of addition by

$$x - y = (a - c, b - d).$$

Fig. 7.11 shows that the point D representing the difference $x - y$ is obtained by translating the directed segment BA along the directed segment BO.

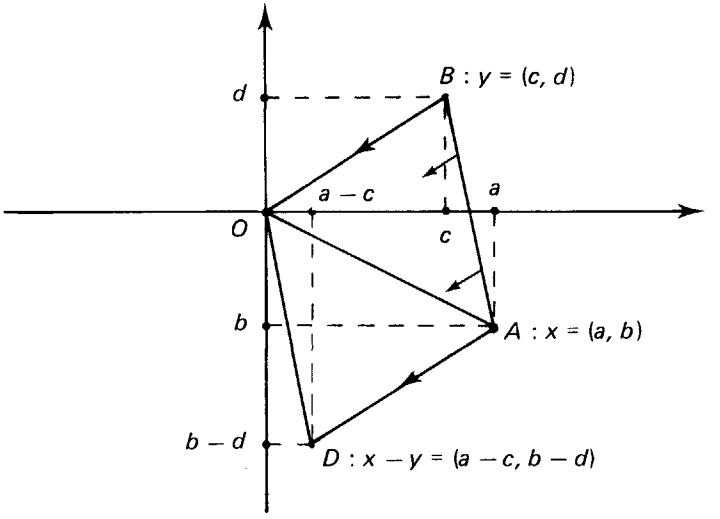


Fig. 7.11

Notice that *ODAB* is now a parallelogram.

Multiplication of complex numbers. The well-known relationship between the Cartesian coordinates and polar coordinates in plane geometry enables us to write, when $x \neq 0$,

$$x = (a, b) = (r \cos \theta, r \sin \theta)$$

where $r = \sqrt{a^2 + b^2}$ is a positive real number and θ is an angle (measured in radians) satisfying $\cos \theta = x/r$ and $\sin \theta = y/r$.

In geometric terms, r is the distance from 0 to x and θ is the angle of rotation from the first coordinate axis to the ray from 0 to x. We shall call r the *modulus* of the complex number x and θ the *argument* (or the *amplitude*) of x. We observe that the modulus r, which is also denoted by $|x|$, is uniquely determined by x, while θ, which is also denoted by $\arg x$, is determinate only in the sense that any multiple of 2π can be added to it. It is therefore convenient to have a *principal value* of the argument and this is subject to an additional condition that $-\pi < \theta \leqslant \pi$ (Fig. 7.12).

The product of two non-zero complex numbers

$$x = (a, b) = (|x| \cos \theta, |x| \sin \theta)$$
$$y = (c, d) = (|y| \cos \psi, |y| \sin \psi)$$

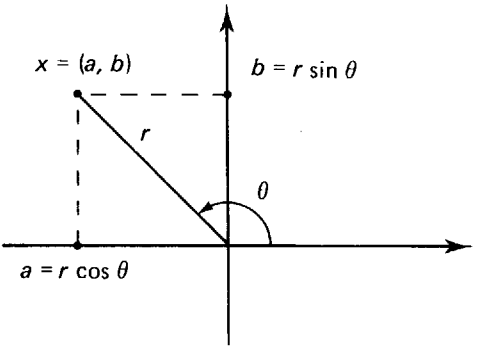


Fig. 7.12

is defined as the complex number

$$d = xy = (|x||y| \cos(\theta + \psi), |x||y| \sin(\theta + \psi)).$$

In other words for the product xy, we have

$$|xy| = |x|\,|y|$$

$$\arg xy = \arg x + \arg y.$$

This is illustrated in Fig. 7.13 below.

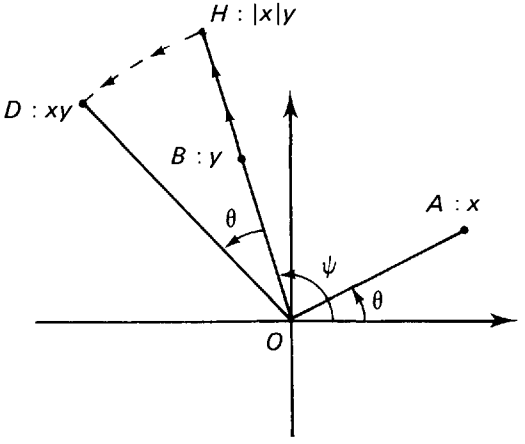


Fig. 7.13

Just like in the procedure used in Sections 7.2 and 7.3, we first prolong (or shorten) OB to OH by a factor $|x| = |OA|$ and then rotate H about O by an angle θ to D. Then H is the point representing the complex number $|x|\,y$ and D is the point representing the product xy.

Using the angle sum formulae of trigonometry,

$$\cos(\theta + \psi) = \cos\theta\cos\psi - \sin\theta\sin\psi$$
$$\sin(\theta + \psi) = \sin\theta\cos\psi + \cos\theta\sin\psi$$

we can also write the definition of product in terms of Cartesian coordinates as

$$xy = (a, b)(c, d) = (ac - bd, ad + bc).$$

The *commutative law* of multiplication

$$xy = yx,$$

the *associative law* of multiplication,

$$x(yz) = (xy)z$$

and the *distributive law*

$$x(y + z) = xy + xz$$

can be easily verified. The complex number $(1, 0)$ has the characteristic property of an neutral element with respect to multiplication,

$$(1, 0)(a, b) = (a, b)(1, 0) = (a, b).$$

If $y \neq 0$, then division by y is also possible and we have

$$\frac{x}{y} = \left(\frac{|x|}{|y|}\cos(\theta - \psi), \frac{|x|}{|y|}\sin(\theta - \psi)\right)$$

or

$$\frac{x}{y} = \frac{(a, b)}{(c, d)} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2}\right).$$

In particular, the reciprocal of a non-zero complex number $x = (a, b) = (r\cos\theta, r\sin\theta)$ is given by

$$x^{-1} = \frac{(1, 0)}{(a, b)} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right)$$
$$= \left(\frac{1}{r}\cos(-\theta), \frac{1}{r}\sin(-\theta)\right).$$

7.4.1 Summary

Complex numbers are ordered pairs of real numbers. The set of all complex numbers is denoted by $\mathbb{C}$. Addition and multiplication of complex numbers are defined by

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \times (c, d) = (ac - bd, ad + bc).$$

The set $\mathbb{C}$ together with the addition and the multiplication defined above constitute the *complex number system* which is also denoted by $\mathbb{C}$.

Complex numbers are represented by points on a plane, which is called the complex plane or the Gaussian plane.

7.5. STANDARD NOTATIONS

Consider the subset R of the set $\mathbb{C}$ of complex numbers that consists of complex numbers of the form $(a, 0)$ where the second coordinate is zero:

$$R = \{(a, 0) \in \mathbb{C} \mid a \in \mathbb{R}\}.$$

In the Gaussian plane, the set R is represented by points of the first coordinate axis. As a set, R is an exact copy of the set $\mathbb{R}$ of real numbers in the sense that every $(a, 0)$ of R corresponds to an element a of $\mathbb{R}$ and vice versa: $(a, 0) \longleftrightarrow a$.

Moreover the set R is closed under addition and multiplication. By this we mean that the sum and product

$$(a, 0) + (b, 0) = (a + b, 0)$$
$$(a, 0) \times (b, 0) = (ab, 0)$$

of any two complex numbers of R are also complex numbers of R. Therefore, we can say that the set R together with addition and multiplication constitute a number system which is a subsystem of the number system $\mathbb{C}$. Furthermore under the one-to-one correspondence $(a, 0) \longleftrightarrow a$ between the sets R and $\mathbb{R}$, the complex number sum and the complex number product correspond to the real number sum and the real number product respectively:

$$(a, 0) + (b, 0) \longleftrightarrow a + b$$
$$(a, 0) \times (b, 0) \longleftrightarrow ab$$

Hence, not only the set R but the number system R itself is an exact copy of the number system $\mathbb{R}$, in the sense that the complex numbers of R behave as if they were real numbers. This means that we may, in the system $\mathbb{C}$, replace the subsystem R, i.e. the set R together with its addition and multiplication, by the system $\mathbb{R}$. After this replacement, the now slightly altered system, which we still denote by $\mathbb{C}$ (though it is now born by the set $(\mathbb{C} \setminus R) \cup \mathbb{R}$), is an extension of the real number system $\mathbb{R}$, in the strict sense of the word. From now on, we may replace every complex number $(a, 0)$ by the real number a. Thus $\mathbb{R}$ is now a subset of $\mathbb{C}$; also addition and multiplication of numbers of $\mathbb{R}$ have the same meaning both

as real numbers and complex numbers. In particular, we have

$$a \times (c, d) = (ac, ad) \quad \text{for every } a \in \mathbb{R} \text{ and } (c, d) \in \mathbb{C}.$$

Now let us consider the second coordinate axis of the Gaussian plane. This consists of complex numbers of the form $(0, b)$, where the first coordinate is zero. Every such complex number is a 'real multiple' of the complex number $(0, 1)$, because

$$(0, b) = b \times (0, 1).$$

If we now denote by i the complex number $(0, 1)$ which has modulus 1 and argument $\pi/2$, we can then write every complex number (a, b) in the new and more convenient form as

$$(a, b) = a + bi$$

because $(a, b) = (a, 0) + (0, b) = a + b \times (0, 1) = a + bi$.

Following the traditional convention, we call the complex number $i = (0, 1)$ the *imaginary unit*. Consequently we also call the first coordinate axis of the Gaussian plane $\mathbb{C}$ the *real axis* and the second coordinate axis the *imaginary axis*. For any complex number

$$x = (a, b) = a + bi$$

we call the real number a the *real part* and the real number b the *imaginary part* of the complex number x and write

$$\mathrm{Re}(x) = a \quad \text{and} \quad \mathrm{Im}(x) = b.$$

We also sometimes call a non-zero complex number with vanishing real part a *purely imaginary number*. Thus bi is a purely imaginary number for every non-zero real number b. It is also customary to call a complex number with non-vanishing imaginary part an *imaginary number*.

Using the new notations, we may write the sum, product and quotient in a more convenient form:

$$(a + bi) + (c + di) = (a + c) + (b + d)\, i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)\, i$$

$$\frac{a + bi}{c + di} = \frac{ac + bd}{c^2 + d^2} - \frac{ad - bc}{c^2 + d^2}\, i$$

In particular,

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}\, i$$

$$\frac{1}{i} = -i$$

We note that the designation of i as the imaginary unit (which was introduced in the last century) is rather antiquated. As we have seen that there is nothing 'imaginary' about the number i and it is just as 'real' as

any real number. The standard notation of writing a complex number in terms of the real unit 1 and the imaginary unit i is a great improvement upon the awkward notation of ordered pair. The benefit of the standard notation will become more apparent when we actually perform calculations with complex numbers in the exercises and in the subsequent sections. We shall see that the standard notation can bring out the essential features of complex numbers more prominently.

Finally a comparison between the number system $\mathbb{R}$ and the extended number system $\mathbb{C}$ will show that there is not very much difference between the two sets of algebraic operations as far as addition, multiplication, subtraction and division are concerned. Naturally, special attention has to be given to the imaginary unit i for which we have

$$ai + bi = (a + b)i, \quad i^2 = ii = -1, \quad -i \cdot i = 1 \quad \text{and} \quad \frac{1}{i} = -i.$$

The counterpart of absolute value $|a|$ of a real number a is the modulus $|x| = \sqrt{a^2 + b^2}$ of a complex number $x = a + bi$, which is also a non-negative real number. Moreover, we take note that we have the natural order relation in the system $\mathbb{R}$, so that given any two real number a and b, we have either $a < b$ or $a = b$ or $a > b$.

There is, however, no such order relation in the system $\mathbb{C}$. Thus it is meaningless to say, for example, one complex number is larger than the other complex number, unless they are both real numbers. Of two arbitrary complex numbers x and y, we can only say that either $x = y$ or $x \neq y$. Thus expressions such as $x \geqslant y$, $y < x$, $x \nleqslant y$ have meaning only when x and y are real numbers; they are, in general, meaningless for complex numbers x and y.

7.6. EXERCISE

Express the following complex numbers in the form $a + bi$, where a and b are both real numbers:

1. $(6 - 2i)(4 + 5i)$
2. $3 + 2i + i\,(5 + i)$
3. $(1 + i)^2$
4. $(2 + i)^3$
5. $(2 - i)^2 - (3 + 2i)^2$
6. $(1 + i)(1 + 2i)(1 + 3i)$

7. $\frac{1}{i}$

8. $(x + yi)(x - yi)$ [x and y are real numbers]

9. $\frac{1}{3 - 4i}$

10. $\frac{1 + i}{1 - i}$

11. $\frac{6 + 5i}{7i}$

12. $\frac{1}{1 + \cos\theta + i \sin\theta}$

Find real numbers x and y satisfying each of the following equations:

13. $\frac{x}{1 + 2i} + \frac{y}{3 + 2i} = \frac{5 + 6i}{8i - 1}$.

14. $x - yi = \frac{1}{2 - 5i}$.

15. $x + yi = (1 - i)^4$

16. Find integers p and q ($0 \leqslant p \leqslant 9$ and $0 \leqslant q \leqslant 9$) such that $(3 + 7i)(p + qi)$ is purely imaginary.

17. If z_1 and z_2 are complex numbers with arguments differing by $\pi/2$, prove that

$$|z_1 - z_2| = |z_1 + z_2|.$$

18. If $\frac{1}{z - ci} = \frac{1}{a + bi} - \frac{1}{a + ci}$ where a, b, c are real numbers, $a \neq 0$ and $b \neq c$, find $|z|^2$.

19. Find the modulus and principal value of the argument of the complex number $-1 + \sqrt{3i}$.

20. If $w = \dfrac{z-1}{z+2}$, where $w = u + vi$ (u and v are real) and $z = \cos\theta + i\sin\theta$, prove that

$$\frac{v}{u} = -3\cot\frac{\theta}{2}$$

and $\quad (u+1)^2 + v^2 = 1.$

21. Solve for z: $\begin{cases} |z| = 1, \\ |z+1| = \sqrt{3}\,|z-1|. \end{cases}$

22. Find the modulus and principal value of the argument of the complex number

$$\frac{1-\cos\theta - i\sin\theta}{1+\cos\theta + i\sin\theta}$$

in each of the following cases: (i) $0 < \theta < \pi$;
(ii) $-\pi < \theta < 0$.

*23. Given that $z^2 = 1 + w^2$, where $z = x + yi$ and $w = a + bi$ (x, y, a, b are real), prove that

(a) $z - w = \dfrac{1}{z+w}$

and hence $\dfrac{x+a}{x-a} = (x+a)^2 + (y+b)^2 = \dfrac{b+y}{b-y}$;

(b) $2x^2 = \sqrt{(1+a^2-b^2)^2 + 4a^2b^2} + 1 + a^2 - b^2$

and

$2y^2 = \sqrt{(1+a^2-b^2)^2 + 4a^2b^2} - 1 - a^2 + b^2.$

7.7. COMPLEX CONJUGATE

Given any complex number $x = a + bi$, the *complex conjugate* (or simply *conjugate*) of x, which is denoted by $\bar{x}$, is defined as the complex number

$$\bar{x} = a - bi.$$

In other words, to obtain the complex conjugate of a complex number, one simply changes the sign of the imaginary part of the complex number. Thus for $x = -3 + 2i$, we get $\bar{x} = -3 - 2i$.

The following formulae involving conjugates can be easily verified:

(a) $\overline{\overline{x}} = x$

(b) $\overline{x+y} = \overline{x} + \overline{y}$; $\overline{xy} = \overline{x}\,\overline{y}$ and $\overline{x/y} = \overline{x}/\overline{y}$

(c) $\operatorname{Re} x = \operatorname{Re} \overline{x} = (x + \overline{x})/2$; $\operatorname{Im} x = -\operatorname{Im} \overline{x} = i(\overline{x} - x)/2$

(d) $x \in \mathbb{R}$ if and only if $x = \overline{x}$.

Recall that the modulus $|x|$ of a complex number $x = a + bi$ is defined as the non-negative real number

$$|x| = \sqrt{a^2 + b^2}$$

and that the argument of a non-zero x, $\arg x = \theta$, is defined by

$$\cos\theta = a/|x| \quad \text{and} \quad \sin\theta = b/|x|$$

Moreover, $x = a + bi$ can be written in polar form (Fig. 7.14) as

$$x = |x|(\cos\theta + i\sin\theta)$$

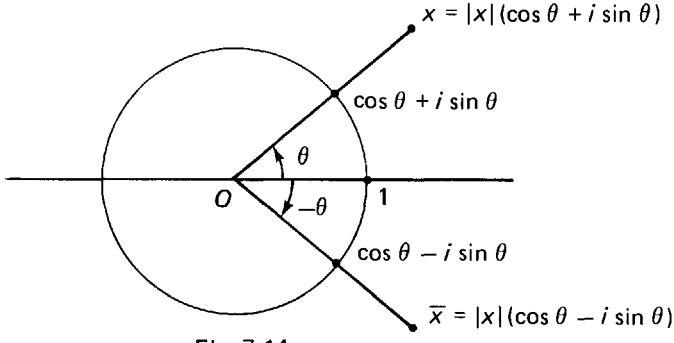


Fig. 7.14

In conjunction with conjugates, we have

(e) $|x| = |\overline{x}| = \sqrt{x\overline{x}}$

(f) $|x| \geqslant \operatorname{Re} x$ and $|x| \geqslant \operatorname{Im} x$.

(g) $1/x = \overline{x}/|x|^2$ if $x \neq 0$

(h) $\arg x = -\arg \overline{x}$ if $x \neq 0$

It follows from (b) and (e) that

$$|xy|^2 = (xy)(\overline{xy}) = (x\overline{x})(y\overline{y}) = |x|^2|y|^2.$$

Therefore we have

(i) $|xy| = |x||y|$.

Similarly, we get

$$\begin{aligned} |x+y|^2 &= (x+y)(\overline{x}+\overline{y}) = x\overline{x} + (x\overline{y} + y\overline{x}) + y\overline{y} \\ &= |x|^2 + 2\operatorname{Re} x\overline{y} + |y|^2 \end{aligned}$$

$$\leqslant |x|^2 + 2|x\bar{y}| + |y|^2$$
$$= |x|^2 + 2|x||y| + |y|^2$$
$$= (|x| + |y|)^2.$$

The equality holds if and only if $\text{Re}\, x\bar{y} = |x\bar{y}|$, i.e. $x\bar{y}$ is real and non-negative. This is precisely the case if and only if $\arg x = \arg y + 2n\pi$ for some integer n. Therefore, we have proved that

(j) $|x + y| \leqslant |x| + |y|$ and the equality sign holds if and only if $\arg x = \arg y + 2n\pi$ for some integer n.

7.8. EQUATIONS WITH REAL COEFFICIENTS

We have set our goal earlier in Section 7.1 to extend the number system $\mathbb{R}$ to a large system $\mathbb{C}$ in which at least the equation $x^2 + 1 = 0$ has a solution. Now that $\mathbb{R}$ has been extended, it is opportune to test the system $\mathbb{C}$ of complex numbers whether it meets the requirement with respect to the solution of the above equation.

Clearly, since $\mathbb{R} \subset \mathbb{C}$, the equation

$$x^2 + 1 = 0$$

is formulated in $\mathbb{C}$ (as well as in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$) in the sense that all its numbers belong to $\mathbb{C}$ and all its symbols have meaning in $\mathbb{C}$. The equation admits no solution in $\mathbb{R}$, but both complex numbers i and $-i$ of $\mathbb{C}$ are solutions of the equation since

$$i^2 + 1 = 0 \quad \text{and} \quad (-i)^2 + 1 = 0.$$

Hence $x^2 + 1 = 0$ admits solutions in $\mathbb{C}$.

We may ask if there are other equations, which admit no solution in $\mathbb{R}$ but become solvable in $\mathbb{C}$. This and other questions will be answered at the end of this section. In the meantime, we shall prove a general theorem on equations with real coefficients and another on quadratic equations with real coefficients.

7.8.1 Theorem. *Let $a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0$ be an equation in the unknown x with real coefficients a_i $(i = 0, 1, \ldots, n)$. If a complex number z is a solution of the equation then the complex conjugate $\bar{z}$ is also a solution of the equation.*

Proof. If z is a solution of the equation

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0,$$

then, after substitution, we get

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0.$$

Taking complex conjugate on both sides, we obtain

$$\begin{aligned} 0 &= \overline{a_n}\,\overline{z^n} + \overline{a_{n-1}}\,\overline{z^{n-1}} + \ldots + \overline{a_1}\,\overline{z} + \overline{a_0} \\ &= a_n \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \ldots + a_1 \bar{z} + a_0 \end{aligned}$$

since $\overline{a_i} = a_i$. Therefore $\bar{z}$ is also a solution of the equation. ∎

The theorem may be rephrased as that the complex solutions of an equation with real coefficients always appear in pairs of conjugates. For example, the pair of conjugates i and $-i$ are solutions of the quadratic equation $x^2 + 1 = 0$. We observe that, since the conjugate of a real number is the number itself, the above statement applies trivially also to the real solutions of such equations.

Consider now the general quadratic equation

$$a^2 x + bx + c = 0$$

with real coefficients a ($\neq 0$), b and c. We know that the equation is solvable in $\mathbb{R}$ if and only if the discriminant

$$D = b^2 - 4ac$$

is non-negative. In this case $\sqrt{D}$ is a real number and

$$ax^2 + bx + c = a\left(x - \frac{-b + \sqrt{D}}{2a}\right)\left(x - \frac{-b - \sqrt{D}}{2a}\right).$$

Thus the solutions of the equation are

$$\frac{-b + \sqrt{D}}{2a} \quad \text{and} \quad \frac{-b - \sqrt{D}}{2a}$$

and the two solutions coincide if $D = 0$. For the remaining case, where $D < 0$, we make use of the real number $\sqrt{-D}$ in the following factorization

$$ax^2 + bx + c = a\left(x - \frac{-b + i\sqrt{-D}}{2a}\right)\left(x - \frac{-b - i\sqrt{-D}}{2a}\right).$$

Here the left-hand side is a quadratic polynomial with real coefficients, while on the right-hand side, we have the product of two linear polynomials with complex coefficients. Therefore the pair of conjugates,

$$\frac{-b + i\sqrt{-D}}{2a} \quad \text{and} \quad \frac{-b - i\sqrt{-D}}{2a}$$

are the solutions of the equation. Thus, in either case and irrespective of the value of the discriminant D, the quadratic equation $ax^2 + bx + c = 0$ admits solutions in $\mathbb{C}$. Using the customary notation below,

$$\sqrt{r} = \begin{cases} \sqrt{r} & \text{if } r \text{ is a real number} \geqslant 0 \\ i\sqrt{-r} & \text{if } r \text{ is a real number} < 0, \end{cases}$$

we may combine both cases in the following theorem.

7.8.2. Theorem. *The solutions of a quadratic equation*

$$ax^2 + bx + c = 0$$

with real coefficients $a (\neq 0)$, b, c *are*

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

7.8.3. *Example.* The solutions of the quadratic equation

$$x^2 + x + 1 = 0$$

are $\omega = \dfrac{-1 + i\sqrt{3}}{2}$ and $\overline{\omega} = \omega^2 = \dfrac{-1 - i\sqrt{3}}{2}$. We may also write

$$\omega = \cos\frac{2\pi}{3} + i \sin\frac{2\pi}{3} \quad \text{and} \quad \overline{\omega} = \cos\frac{4\pi}{3} + i \sin\frac{4\pi}{3}.$$

We have now seen that besides $x^2 + 1 = 0$, all quadratic equations with real coefficients are solvable in $\mathbb{C}$. We may therefore ask the following question on the provision of solutions of polynomial equations with coefficients in $\mathbb{R}$.

7.8.4. *Question.* Are all polynomial equations in one unknown with real coefficients solvable in $\mathbb{C}$?

According to the discussion in Section 7.1, if the answer to this question is negative, then we would have to find one such equation that is unsolvable in $\mathbb{C}$ and try to extend the system $\mathbb{C}$ to some still larger system to accommodate its solutions. On the other hand, if the answer to the above question is affirmative, then the inadequancy in the provision of solutions of equations, which are formulated in $\mathbb{R}$, is entirely removed. But there are more equations formulated in $\mathbb{C}$ than in $\mathbb{R}$. Therefore we should now examine also the provision of solutions in $\mathbb{C}$ of polynomial equations with complex coefficients. This means that we have to answer the following question.

7.8.5. *Question.* Are all polynomial equations in one unknown with complex coefficients solvable in $\mathbb{C}$?

It turns out that the answers to both questions are affirmative and they are formulated in the so-called *fundamental theorem of algebra.*

7.8.6. Fundamental theorem of algebra. *Let $f(x)$ be a polynomial of degree $n > 0$ with complex coefficients. Then the equation $f(x) = 0$ in the unknown x has at least one and at most n distinct solutions in $\mathbb{C}$.*

The first proof of this theorem is given by C.F. Gauss. Since then many new proofs have been found. Unfortunately all such proofs are well above the level of this course. Also since we shall have few opportunities to apply this theorem in the subsequent sections, the reader is therefore asked to note that by virtue of this theorem, it is not necessary to further extend the system $\mathbb{C}$. Thus the number system $\mathbb{C}$ is completely adequate as far as the provision of solutions of equations is concerned.

There is however another point of view regarding the possible extension of the number system $\mathbb{C}$. We have accepted the geometric interpretation of the real number system $\mathbb{R}$ as a one-dimensional number system and the complex number system $\mathbb{C}$ as a two-dimensional number system. We may therefore ask if it is possible and useful to extend the system $\mathbb{C}$ to some still higher dimensional number system. In order not to disrupt the development of our main theme of study, we shall leave this question for the time being and return for it in the Appendix 7.18 to this chapter.

7.9. EXERCISE

1. If $|z_1| = |z_2|$ but $z_1 \neq z_2$ and $z_1 + z_2 \neq 0$, prove that $\dfrac{z_1 + z_2}{z_1 - z_2}$ is purely imaginary.

2. Given that $|z + ai| = |z + bi|$ where a and b are distinct real numbers, prove that
$$z - \bar{z} = -(a + b)i.$$

3. Given that $4|z + 1| = |z + 16|$, deduce that $|z| = 4$.

4. (a) Using the inequality $|z_1 + z_2| \leqslant |z_1| + |z_2|$ or otherwise, prove that $|z_1 - z_2| \geqslant |z_1| - |z_2|$ for any two complex numbers z_1 and z_2.
   (b) Show that the equation $z^4 + z + 2 = 0$ cannot have a root with modulus less than 1.

5. Prove that $2|z_1|^2 + 2|z_2|^2 = |z_1 - z_2|^2 + |z_1 + z_2|^2$ for any two complex numbers z_1 and z_2.

6. (a) Let a, b, c and d be any four complex numbers, prove that
$$|(a-c)(b-d)| \leqslant |(a-d)(b-c)| + |(c-d)(a-b)|.$$
(b) If A, B, C and D are any four points in a plane, prove that
$$AC \cdot BD \leqslant AB \cdot CD + AD \cdot BC.$$

Solve the following quadratic equations, expressing your answers in the form $a + bi$, where a and b are real:

7. $x^2 + 4x + 13 = 0$.

8. $4x^2 + 9 = 0$.

9. $x^2 - 2x \cos\theta + 1 = 0$.

10. $3x^2 - (2 + 11i)x + 3 - 5i = 0$.

11. Find a quadratic equation whose roots are $2 + \sqrt{3i}$ and $2 - \sqrt{3i}$.

*12. Find the condition for one root of the equation
$$z^2 + 2(a + bi)z + (c + di) = 0$$
where a, b, c and d are real numbers, to be real.

*13. If z_1 and z_2 are the roots of the quadratic equation
$$az^2 + bz + c = 0$$
with a, b, c real and $b^2 < 4ac$, show that
$$\left|\frac{z_1}{z_2}\right| = 1$$
and $$\arg\left(\frac{z_1}{z_2}\right) = 2\cos^{-1}\sqrt{\frac{b^2}{4ac}}$$

7.10. DE MOIVRE'S THEOREM

Complex numbers of modulus 1 are represented by points on the unit circle of the Gaussian plane. They have the interesting property that their powers remain on the unit circle. In polar form
$$z = \cos\theta + i\sin\theta$$

when raised to a power by a positive exponent n, it becomes

$$z^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

by the product formula of Section 7.4.

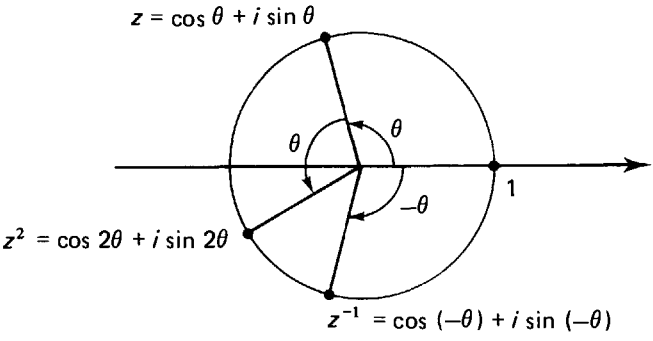


Fig. 7.15

Thus, if we rotate the point 1 about 0 by θ, we get the point z, by 2θ the point z^2, by 3θ the point z^3, etc (Fig. 7.15).

The last equation also holds true for negative integers $-n$ $(n > 0)$, for

$$\begin{aligned} z^{-n} &= \frac{1}{(\cos\theta + i\sin\theta)^n} = \frac{1}{\cos n\theta + i\sin n\theta} \\ &= \frac{\cos n\theta - i\sin n\theta}{\cos^2 n\theta - i^2\sin^2 n\theta} = \cos n\theta - i\sin n\theta \\ &= \cos(-n\theta) + i\sin(-n\theta). \end{aligned}$$

Therefore we have proved the following theorem which is named after the French mathematican, Abraham De Moivre (1667-1754).

7.10.1. Theorem (De Moivre). *For any integer n,*

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

Expanding the left-hand side of the equation by the binomial theorem, and equating respectively the real parts and the imaginary parts of both sides, we obtain the following formulae.

7.10.2. De Moivre's Formulae. *For any positive integer n,*

$$\cos n\theta = \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta - \ldots$$

$$\sin n\theta = n\cos^{n-1}\theta\sin\theta - \binom{n}{3}\cos^{n-3}\theta\sin^3\theta + \ldots$$

These two formulae express $\cos n\theta$ and $\sin n\theta$ as polynomials in $\cos\theta$ and $\sin\theta$ with integer coefficients. We observe here that they are expressions of relations among real numbers. We may well imagine that they can be proved by methods of trigonometry. Such a proof will involve induction and the angle sum formula of trigonometric functions. Undoubtedly, it will be more complicated than the derivation from De Moivre's theorem. The derivation can serve as an example in which complex numbers may be put to work to give results concerning real numbers. We illustrate the derivation of the formulae for $n = 5$ in the following example.

7.10.3. *Example.* Expanding $(\cos\theta + i\sin\theta)^5$ by the binomial theorem, we get

$$\begin{aligned}\cos 5\theta + i\sin 5\theta &= (\cos\theta + i\sin\theta)^5\\ &= \cos^5\theta + 5i\cos^4\theta\sin\theta + 10i^2\cos^3\theta\sin^2\theta +\\ &\quad 10i^3\cos^2\theta\sin^3\theta + 5i^4\cos\theta\sin^4\theta + i^5\sin^5\theta\\ &= (\cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta) +\\ &\quad i(5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta).\end{aligned}$$

Thus comparing real and imaginary parts on both ends, we get

$$\cos 5\theta = \cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta$$
$$\sin 5\theta = 5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta.$$

We can further derive from these two expressions that

$$\begin{aligned}\cos 5\theta &= \cos^5\theta - 10\cos^3\theta(1 - \cos^2\theta) + 5\cos\theta(1 - \cos^2\theta)^2\\ &= 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta\end{aligned}$$

$$\begin{aligned}\sin 5\theta &= \sin^5\theta - 10(1 - \sin^2\theta)\sin^3\theta + 5(1 - \sin^2\theta)^2\sin\theta\\ &= 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta.\end{aligned}$$

The two parallel formulae for $\cos 5\theta$ and $\sin 5\theta$ that we have just obtained in the last example can be further exploited to get other classical results in trigonometry.

7.10.4. *Example.* Evaluate $\sin\pi/5$.

Solution. Take the last formula

$$\sin 5\theta = 16\sin^5\theta - 20\sin^3\theta + 5\sin\theta.$$

Substituting $\theta = \pi/5$ and denoting $\sin\pi/5$ by x, we get a quintic equation (i.e. an equation of degree 5)

$$16x^5 - 20x^3 + 5x = 0$$

whose roots are $\sin 0, \sin \pi/5, \sin 2\pi/5, \sin -\pi/5$ and $\sin -2\pi/5$, since these values satisfy the original trigonometric equation. Cancelling the factor x in the last equation, we have a quartic equation (i.e. an equation of degree 4)

$$16x^4 - 20x^2 + 5 = 0$$

with roots $\sin(\pm\pi/5)$ and $\sin(\pm 2\pi/5)$. Therefore all that remains to be done is to find the roots of this quartic equation and identify $\sin \pi/5$ among them. Regarding it as a quadratic equation in x^2, we solve for x^2 to get

$$x^2 = \frac{10 \pm 2\sqrt{5}}{16}.$$

Therefore, the four roots of the quartic equation are:

$$s_1 = \frac{1}{4}\sqrt{10 + 2\sqrt{5}} \qquad s_2 = \frac{1}{4}\sqrt{10 - 2\sqrt{5}}$$

$$s_3 = -\frac{1}{4}\sqrt{10 - 2\sqrt{5}} \qquad s_4 = -\frac{1}{4}\sqrt{10 + 2\sqrt{5}}$$

Since $$\sin\frac{-2\pi}{5} < \sin\frac{-\pi}{5} < 0 < \sin\frac{\pi}{5} < \sin\frac{2\pi}{5}$$

$$s_4 < s_3 < 0 < s_2 < s_1.$$

We get the desired value

$$\sin\frac{\pi}{5} = \frac{1}{4}\sqrt{10 - 2\sqrt{5}}.$$ ■

Another application of De Moivre's Theorem is to express powers of $\sin\theta$ and $\cos\theta$ in terms of sines and cosines of multiples of θ. Consider the complex numbers

$$z = \cos\theta + i\sin\theta \quad \text{and} \quad \frac{1}{z} = \cos\theta - i\sin\theta.$$

By De Moivre's Theorem,

$$z^n = \cos n\theta + i\sin n\theta \quad \text{and} \quad \frac{1}{z^n} = \cos n\theta - i\sin n\theta.$$

Therefore

$$\cos n\theta = \frac{1}{2}\left(z^n + \frac{1}{z^n}\right) \quad \text{and} \quad \sin n\theta = \frac{1}{2i}\left(z^n - \frac{1}{z^n}\right).$$

Let us now use these two formulae to obtain an expression for $\sin^7\theta$.

7.10.5. *Example.* Express $\sin^7\theta$ in terms of sines of multiples of θ.

Solution. Let $z = \cos\theta + i\sin\theta$. Then

$$(2i\sin\theta)^7 = (z - \frac{1}{z})^7$$

$$= z^7 - 7z^5 + 21z^3 - 35z + \frac{35}{z} - \frac{21}{z^3} + \frac{7}{z^5} - \frac{1}{z^7}$$

$$= (z^7 - \frac{1}{z^7}) - 7(z^5 - \frac{1}{z^5}) + 21(z^3 - \frac{1}{z^3}) - 35(z - \frac{1}{z})$$

$$= 2i(\sin 7\theta - 7\sin 5\theta + 21\sin 3\theta - 35\sin\theta).$$

Therefore $\sin^7\theta = \frac{1}{64}(-\sin 7\theta + 7\sin 5\theta - 21\sin 3\theta + 35\sin\theta)$. ∎

Clearly the method used in the above example is very much superior to those of elementary trigonometry. This shows once more the advantage of putting complex numbers to work for us. A further example along the same lines is to obtain sums of finite series of trigonometric terms.

7.10.6. *Example.* Prove that for all values of θ, which are not multiples of 2π,

$$1 + \cos\theta + \cos 2\theta + \ldots + \cos n\theta = \frac{1}{\sin\frac{\theta}{2}}(\sin\frac{n+1}{2}\theta\cos\frac{n}{2}\theta)$$

$$\sin\theta + \sin 2\theta + \ldots + \sin n\theta = \frac{1}{\sin\frac{\theta}{2}}(\sin\frac{n+1}{2}\theta\sin\frac{n}{2}\theta).$$

Proof. Consider the well-known identity

$$(1 - z)(1 + z + z^2 + \ldots + z^n) = 1 - z^{n+1}$$

which holds for real numbers z as well as for complex numbers z. Let θ be different from multiples of 2π, and $z = \cos\theta + i\sin\theta$. Then $z \neq 1$ and

$$1 + z + z^2 + \ldots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

Substituting $z = \cos\theta + i\sin\theta$ and simplifying by De Moivre's Theorem, we get

$$(1 + \cos\theta + \cos 2\theta + \ldots + \cos n\theta) + i(\sin\theta + \sin 2\theta + \ldots + \sin n\theta)$$

$$= \frac{1 - \cos(n+1)\theta - i\sin(n+1)\theta}{1 - \cos\theta - i\sin\theta}$$

$$= \frac{2\sin^2\frac{n+1}{2}\theta - i\,2\sin\frac{n+1}{2}\theta\cos\frac{n+1}{2}\theta}{2\sin^2\frac{\theta}{2} - i\,2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}$$

$$= \left(\frac{\sin\frac{n+1}{2}\theta}{\sin\frac{\theta}{2}}\right)\left(\frac{\sin\frac{n+1}{2}\theta - i\cos\frac{n+1}{2}\theta}{\sin\frac{\theta}{2} - i\cos\frac{\theta}{2}}\right)\left(\frac{i}{i}\right)$$

$$= \left(\frac{\sin\frac{n+1}{2}\theta}{\sin\frac{\theta}{2}}\right)\left(\frac{\cos\frac{n+1}{2}\theta + i\sin\frac{n+1}{2}\theta}{\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}}\right)$$

$$= \left(\frac{\sin\frac{n+1}{2}\theta}{\sin\frac{\theta}{2}}\right)\left(\cos\frac{n}{2}\theta + i\sin\frac{n}{2}\theta\right)$$

Therefore,

$$1 + \cos\theta + \cos 2\theta + \ldots + \cos n\theta = \frac{1}{\sin\frac{\theta}{2}}\sin\frac{n+1}{2}\theta\cos\frac{n}{2}\theta$$

$$\sin\theta + \sin 2\theta + \ldots + \sin n\theta = \frac{1}{\sin\frac{\theta}{2}}\sin\frac{n+1}{2}\theta\sin\frac{n}{2}\theta.$$

7.11. EXERCISE

1. If $z = \cos\theta + i\sin\theta$, prove that
$$\frac{1 - \frac{1}{z^2}}{1 + \frac{1}{z^2}} = i\tan\theta.$$

2. Show that, for any positive integer n and real number θ,
$$(\cos\theta - i\sin\theta)^n = \cos n\theta - i\sin n\theta.$$

3. If n is an integer, show that
$$(1 + i\tan\theta)^n + (1 - i\tan\theta)^n = 2\sec^n\theta\cos n\theta.$$

4. Show that, for any positive integer n and real number θ,
$$\Sigma_{r=0}^{n}\binom{n}{r}\cos r\theta = 2^n\cos^n\frac{\theta}{2}\cos\frac{n\theta}{2}.$$

5. Show that, for any positive integer n,
$$\Sigma_{r=0}^{n}\binom{n}{r}\sin(n-r)\theta\sin^{n-r}\varphi\sin^r(\theta - \varphi)$$

$= \sin^n \theta \sin n\varphi$

and $\Sigma_{r=0}^{n} \binom{n}{r} \cos(n-r)\theta \sin^{n-r} \varphi \sin^r(\theta - \varphi)$

$= \sin^n \theta \cos n\varphi$

for any positive integer n.

(Hint: consider $z = \cos\theta + i \sin\theta$.)

6. (a) Prove that $\cos 5\theta = 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta$
and $\cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1$.

*(b) Deduce that $\cos(2\pi/9)$, $\cos(4\pi/9)$ and $\cos(8\pi/9)$ are the three roots of the equation $8x^3 - 6x + 1 = 0$.

7. (a) Show that $\cot 7\theta = \dfrac{1 - 21\tan^2\theta + 35\tan^4\theta - 7\tan^6\theta}{7\tan\theta - 35\tan^3\theta + 21\tan^5\theta - \tan^7\theta}$.

*(b) Prove that $\tan^2(\pi/14)$, $\tan^2(3\pi/14)$ and $\tan^2(5\pi/14)$ are the three roots of the equation $7x^3 - 35x^2 + 21x - 1 = 0$.

(c) Deduce that $\prod_{k=0}^{2} \tan\dfrac{(2k+1)\pi}{14} = \dfrac{1}{\sqrt{7}}$.

8. Let $C_n = 1 + \Sigma_{k=1}^{n} x^k \cos k\theta$ and $S_n = \Sigma_{k=1}^{n} x^k \sin k\theta$
where x is real and θ is not a multiple of π.

(a) Show that $C_n = \dfrac{1 - x\cos\theta - x^{n+1}\cos(n+1)\theta + x^{n+2}\cos n\theta}{1 - 2x\cos\theta + x^2}$

and $S_n = \dfrac{x\sin\theta - x^{n+1}\sin(n+1)\theta + x^{n+2}\sin n\theta}{1 - 2x\cos\theta + x^2}$

(b) If $|x| < 1$, find $\lim_{n\to\infty} C_n$ and $\lim_{n\to\infty} S_n$.

(c) Prove that $\Sigma_{k=0}^{\infty} \cos^k\theta \sin(\alpha + k\theta) = \operatorname{cosec}\theta \cos(\theta - \alpha)$.

9. Let $C = \Sigma_{k=0}^{n} (-1)^k \binom{n}{k} \cos 2k\theta$

and $S = \Sigma_{k=1}^{n} (-1)^k \binom{n}{k} \sin 2k\theta$.

Show that $C = \begin{cases} (-1)^{n/2} (2\sin\theta)^n \cos n\theta & \text{if } n \text{ is even} \\ (-1)^{(n-1)/2} (2\sin\theta)^n \sin n\theta & \text{if } n \text{ is odd} \end{cases}$

and $S = \begin{cases} (-1)^{n/2} (2\sin\theta)^n \sin n\theta & \text{if } n \text{ is even} \\ (-1)^{(n-1)/2} (2\sin\theta)^n \cos n\theta & \text{if } n \text{ is odd.} \end{cases}$

(Hint: consider $(1 - z^2)^n$ where $z = \cos\theta + i\sin\theta$.)

*10. Let $z_1 = \cos A + i \sin A$,
$z_2 = \cos B + i \sin B$
and $z_3 = \cos C + i \sin C$
such that A, B and C are the angles of a triangle.
(a) Show that $z_1 z_2 z_3 = -1$.
(b) Prove that $\cos^2 A + \cos^2 B + \cos^2 C = 1 - 2\cos A \cos B \cos C$
and $\sin 2A + \sin 2B + \sin 2C = 4\sin A \sin B \sin C$.
(Hint: consider the products $(z_1 + \frac{1}{z_1})(z_2 + \frac{1}{z_2})(z_3 + \frac{1}{z_3})$
and $(z_1 - \frac{1}{z_1})(z_2 - \frac{1}{z_2})(z_3 - \frac{1}{z_3})$.

*11. (a) Show that $\sin 7\theta = -\sin\theta\,(64\sin^6\theta - 112\sin^4\theta + 56\sin^2\theta - 7)$.
(b) Deduce that $\sin^2(\pi/7)$, $\sin^2(2\pi/7)$ and $\sin^2(3\pi/7)$ are the three roots of the equation
$$64x^3 - 112x^2 + 56x - 7 = 0$$
and that $\sec(2\pi/7)$, $\sec(4\pi/7)$ and $\sec(6\pi/7)$ are the three roots of the equation
$$x^3 + 4x^2 - 4x - 8 = 0.$$
(c) Show that (i) $\Sigma_{k=1}^{3} \sin^2(\frac{k\pi}{7}) = \frac{7}{4}$,
(ii) $\Sigma_{k=1}^{3} \operatorname{cosec}^2(\frac{k\pi}{7}) = 8$
and (iii) $\Sigma_{k=1}^{3} \sec\frac{2k\pi}{7} = -4$.

7.12. THE *n*-th ROOTS

Given a complex number z and a positive integer n, we say that a number x is an *n-th root* of z if
$$x^n = z.$$
Regarding this as an equation of degree n in the unknown x, we may conclude, by the fundamental theorem of algebra 7.10.4, that z has at least one and at most n distinct n-th roots. However, it is not necessary to use such powerful theorem to come to an even more precise conclusion.

7.12.1. Theorem. *Every complex number $z \neq 0$ has exactly n distinct n-th roots.*

Proof. Writing in the polar form

$$z = r(\cos\theta + i\sin\theta)$$
$$x = t(\cos\psi + i\sin\psi)$$

we can express the equation

$$x^n = z$$

as

$$t^n(\cos n\psi + i\sin n\psi) = r(\cos\theta + i\sin\theta).$$

Since any two complex numbers are equal if and only if they have the same modulus and the same argument (up to a multiple of 2π), it follows that a necessary and sufficient condition for x to be an n-th root of z is that

$$t = \sqrt[n]{r} \quad \text{and} \quad n\psi = \theta + 2h\pi$$

for some integer h. Therefore we have just one possibility for the modulus of x, which is

$$t = \sqrt[n]{r}$$

while we have for the argument ψ of x among others the n possibilities

$$\frac{\theta}{n}, \frac{\theta + 2\pi}{n}, \frac{\theta + 4\pi}{n}, \ldots, \frac{\theta + 2(n-1)\pi}{n}$$

The difference between any two of these n values is less than 2π; therefore they are the arguments of n distinct complex numbers. Thus

$$x_j = \sqrt[n]{r}\left(\cos\frac{\theta + 2j\pi}{n} + i\sin\frac{\theta + 2j\pi}{n}\right), \quad j = 0, 1, \ldots, n-1$$

are n distinct n-th roots of $z = r(\cos\theta + i\sin\theta)$.

On the other hand if

$$y = \sqrt[n]{r}\left(\cos\frac{\theta + 2h\pi}{n} + i\sin\frac{\theta + 2h\pi}{n}\right)$$

is any n-th root of z, then

$$h \equiv j \pmod{n}$$

for exactly one value of $j = 0, 1, 2, \ldots, n-1$. Hence $\arg y = (\theta + 2h\pi)/n$ differs from $\arg x_j = (\theta + 2j\pi)/n$ by a multiple of 2π and $y = x_j$. Therefore we have exactly n distinct n-th roots $x_0, x_1, \ldots, x_{n-1}$ of the non-zero complex number z. ∎

On the Gaussian plane, the n-th roots $x_0, x_1, \ldots x_{n-1}$ of $z \neq 0$ are all points on a circle of radius $\sqrt{|z|}$ centred at 0. Fig. 7.16 illustrates the four 4th roots of z for $n = 4$.

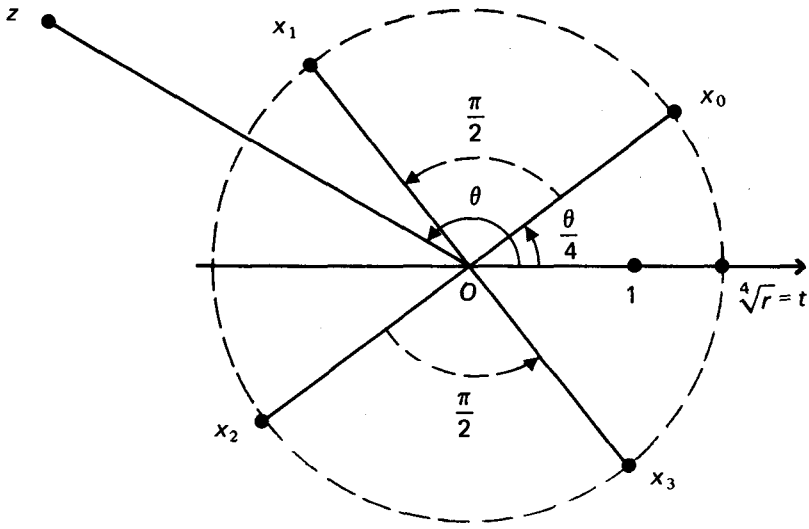


Fig. 7.16

In particular the four 4th roots of -1 (Fig. 7.17) are:

$$x_0 = \cos\frac{\pi}{4} + i\sin\frac{\pi}{4} = \frac{1}{2}(\sqrt{2} + i\sqrt{2})$$

$$x_1 = \cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4} = \frac{1}{2}(-\sqrt{2} + i\sqrt{2})$$

$$x_2 = \cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4} = \frac{1}{2}(-\sqrt{2} - i\sqrt{2})$$

$$x_3 = \cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4} = \frac{1}{2}(\sqrt{2} - i\sqrt{2}).$$

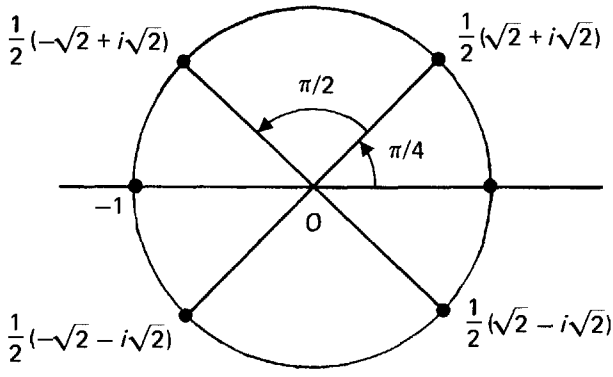


Fig. 7.17

The n-th roots of $z = 1$ are called the *n-th roots of unity.* By the results of Theorem 7.14.1, these are the complex numbers

$$x_0 = 1, \quad x_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \quad x_2 = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n}, \ldots,$$

$$x_{n-1} = \cos\frac{2(n-1)\pi}{n} + i\sin\frac{2(n-1)\pi}{n}.$$

It is easily recognized that all these are powers of n-th root of unity,

$$\xi = x_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}.$$

On the Gaussian plane, they are the points on the unit circle with arguments being multiples of $2\pi/n$. They are the vertices of a regular polygon of n sides (Fig. 7.18).

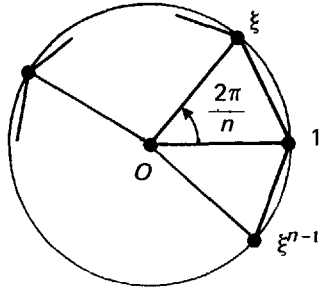


Fig. 7.18

In other words,

$$\xi, \xi^2, \xi^3, \ldots, \xi^n$$

are all the n-th roots of unit. We call a complex number x a *primitive n-th root of unity* if each n-th root of unit is a power of x. Thus ξ is a primitive n-th root of unity. So are $(\sqrt{2} + i\sqrt{2})/2$, $(-\sqrt{2} + i\sqrt{2})/2$, $(-\sqrt{2} - i\sqrt{2})/2$ and $(\sqrt{2} - i\sqrt{2})/2$ primitive 8th roots of unity, but the other four 8th roots of unity 1, i, -1 and $-i$ are not primitive.

For $n = 3$, the three cube roots of unity (Fig. 7.19) are

$$1 = \cos\frac{0}{3} + i\sin\frac{0}{3}$$

$$\omega = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = \frac{1}{2}(-1 + i\sqrt{3})$$

$$\omega^2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = \frac{1}{2}(-1 - i\sqrt{3}).$$

Both ω and ω^2 are primitive cube roots of unity while 1 is obviously not.

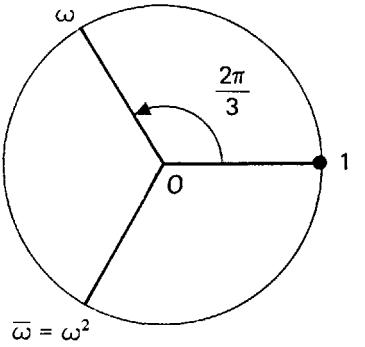


Fig. 7.19

Obviously primitive n-th roots of unity are of special interest because they can generate all other n-th roots of unity. In general, if

$$\xi = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$$

is the n-th root of unity with the least positive argument, then we can show with results in Chapter 4 that ξ^m is primitive if and only if n and m are relatively prime.

7.13. EXERCISE

1. Find the square roots of i.

2. Express the square roots of $5 + 2i$ in the form $a + bi$ where a and b are integers.

3. Find the cube roots of -1.

4. Find the cube roots of $1 - \cos\theta - i\sin\theta$ where $0 < \theta < 2\pi$.

5. Resolve $x^6 + 1$ into the product of real quadratic factors.

6. (a) Solve the equation $x^7 + 1 = 0$.
   (b) Show that $\cos\frac{\pi}{7} + \cos\frac{3\pi}{7} + \cos\frac{5\pi}{7} = \frac{1}{2}$.

7. Let ω $(\neq 1)$ be an n-th root of 1, prove that $\sum_{k=0}^{n-1} \omega^k = 0$.

8. If ω is a primitive n-th root of 1, prove that
   (a) $\prod_{k=1}^{n-1} (1 - \omega^k) = n$
   (b) $\sum_{k=0}^{n-1} \omega^{kr} = \begin{cases} 0 & \text{if } r \text{ is not a multiple of } n \\ n & \text{if } r \text{ is a multiple of } n \end{cases}$

9. (a) Prove that $z^{13} - 1 = (z - 1)\prod_{k=1}^{6} (z^2 - 2z\cos\frac{2k\pi}{13} + 1)$.
   (b) Deduce that
   $(\cos\frac{2\pi}{13} + \cos\frac{10\pi}{13})(\cos\frac{4\pi}{13} + \cos\frac{6\pi}{13})(\cos\frac{8\pi}{13} + \cos\frac{12\pi}{13}) = \frac{1}{8}$.

10. (a) Solve the equation $(z + 1)^8 - z^8 = 0$.
    (b) Prove that $(z + 1)^8 - z^8 = \frac{1}{16}(2z + 1)\prod_{k=1}^{3} [4z^2 + 4z + \operatorname{cosec}^2(\frac{k\pi}{8})]$.
    Hence show that
    $$16(\cos^{16}\theta - \sin^{16}\theta) = \cos 2\theta \prod_{k=1}^{3} [\cos^2 2\theta + \cot^2(\frac{k\pi}{8})].$$

11. Solve the equation $z^n + (z - a)^n = 0$ where a is a non-zero real number and n is a positive integer.

*12. If ω is a primitive 9th root of 1, obtain the value of

$$\sum_{k=1}^{4} (2 + \omega^{-k} + \omega^k)^2.$$

Deduce that $\sum_{k=1}^{4} \cos^4 (\frac{k\pi}{9}) = \frac{19}{16}$.

*13. (a) Show that $z^{2n} - 1 = (z - 1)(z + 1) \prod_{k=1}^{n-1} (z^2 - 2z \cos \frac{k\pi}{n} + 1)$ where n is any positive integer greater than 1.

(b) Deduce that $\prod_{k=1}^{n-1} (\cos \theta - \cos \frac{k\pi}{n}) = \frac{\sin n\theta}{2^{n-1} \sin \theta}$ if $\sin \theta \neq 0$.

*14. Let n be a positive integer.

(a) Find the $2n$ distinct complex roots of the equation

$$(1 + z)^{2n} + (1 - z)^{2n} = 0.$$

Express the roots in the form $a + bi$ where a and b are real numbers.

(b) Show that $(1 + z)^{2n} + (1 - z)^{2n} = 2 \prod_{k=0}^{n-1} [z^2 + \tan^2 \frac{(2k + 1)\pi}{4n}]$.

(c) Show that $\sum_{k=0}^{n-1} \sec^2 [\frac{(2k + 1)\pi}{4n}] = 2n^2$.

*15. If n is a positive integer, show that

$$z^{2n} - 2z^n \cos \alpha + 1 = \prod_{k=0}^{n-1} (z^2 - 2z \cos \frac{\alpha + 2k\pi}{n} + 1).$$

Deduce that

(i) $\cos n\theta - \cos n\varphi = 2^{n-1} \prod_{k=0}^{n-1} [\cos \theta - \cos (\varphi + \frac{2k\pi}{n})]$;

(ii) $\sin n\alpha = 2^{n-1} \prod_{k=0}^{n-1} \sin (\alpha + \frac{k\pi}{n})$;

(iii) $\cos n\alpha = 2^{n-1} \prod_{m=0}^{n-1} \sin [\alpha + \frac{(2m + 1)\pi}{2n}]$ where $0 < \alpha < \frac{\pi}{n}$.

7.14. GEOMETRY OF COMPLEX NUMBERS

The central idea of plane analytic geometry is the establishment of a

correspondence between pairs of real numbers and points in the plane, thereby making possible a correspondence between curves in the plane and equations in *two real variables*. Under this correspondence, for each curve in the plane, there is a definite equation $f(x, y) = 0$, and for each equation, there is a definite curve. Moreover, the algebraic properties of the equation correspond to the geometric properties of the associated curve, so that the often difficult task of proving a theorem in geometry is shifted to that of proving a corresponding theorem in algebra. The correspondence between complex numbers and points in the plane can be similarly used for the development of a geometry of the complex plane. Here a point on the plane corresponds to a single complex number; therefore curves in the plane correspond to equations of a *single complex variable*. Besides the reduction of one variable on the algebraic side of the interplay of geometry and algebra, we shall also have the very rich algebraic properties of the complex numbers at our disposal.

We shall adopt the convention of analytic geometry and use the notation $P(x)$ to express the fact that the point P on the plane and the complex number $x = a + bi$ are associated in the sense that the Cartesian coordinates of P are (a, b). Sometimes, when no confusion will arise, we may just refer to x as a point meaning that it is the point in the plane associated to x in the above sense.

Let us first rewrite some well-known formulae of analytic geometry in terms of complex numbers. The so-called distance formula is easily given in terms of modulus. Given two points $P(x)$ and $Q(y)$ on the plane, the distance between P and Q is given by

$$|x - y| = \sqrt{(a - c)^2 + (b - d)^2}$$

where $x = a + bi$ and $y = c + di$. Using this formula, we find that for two points $A(s)$ and $B(t)$, the perpendicular bisector of the segment AB is given by the equation

$$|z - s| = |z - t|$$

in the complex variable z.

Obviously, the formula for angle measurement is given in terms of argument. Let $P(x)$ and $Q(y)$ be two distinct points on the plane. If R is the point that represents the complex number $y - x : R(y - x)$, then the rays PQ and OR will have the same direction (Fig. 7.20).

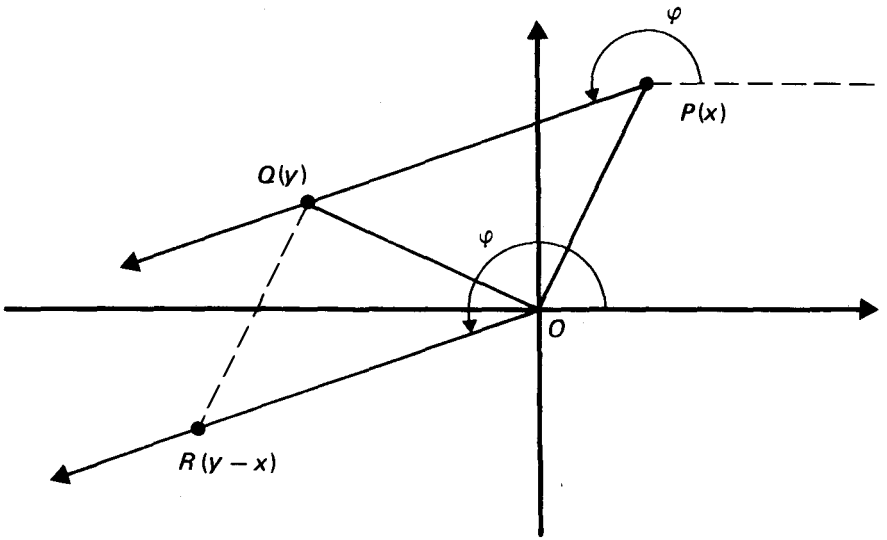


Fig. 7.20

Therefore the angle φ between the positive real axis and the ray PQ is identical to that between the positive real axis and the ray OR. Since the latter is given by $\arg(y - x)$, so $\varphi = \arg(y - x)$ (Fig. 7.21).

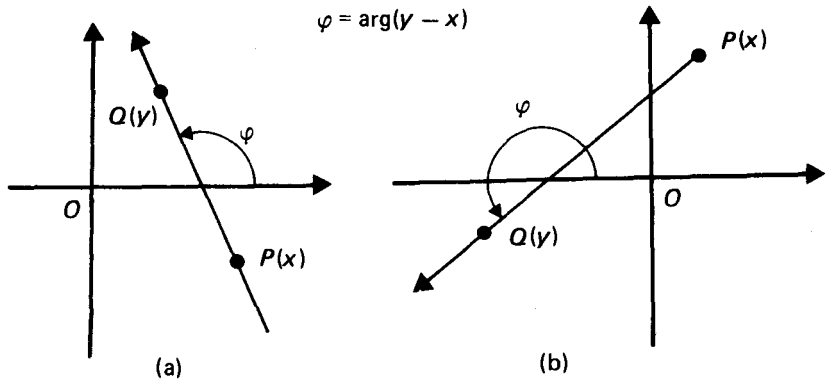


Fig. 7.21(a) & (b)

Let $P(x)$, $Q(y)$ and $S(z)$ be three distinct points. If we measure the angle $\psi = \angle SPQ$ by the angle of rotation that takes the ray PS to the ray PQ, then using the positive real axis as a fixed reference and the formula above (Fig. 7.22), we get

$$\psi = \arg(y - x) - \arg(z - x) = \arg\left(\frac{y - x}{z - x}\right).$$

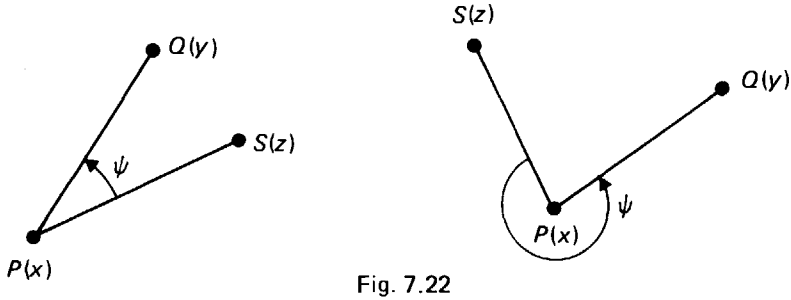


Fig. 7.22

This very neat and tidy formula for angle measurement is clearly a great improvement upon the very cumbersome expression in terms of Cartesian coordinates. As an application of this formula, we study the following example.

7.14.1. *Example.* Let $A_1(x_1)$, $A_2(x_2)$ and $A_3(x_3)$ be three distinct points on a circle. A necessary and sufficient condition for a point $P(z)$ to lie on the circular arc $A_1\ A_2\ A_3$ is that

$$\angle A_1A_2A_3 = \angle A_1PA_3$$

By the above angle formula, this becomes an equation in the unknown z,

$$\arg\left(\frac{x_3 - x_2}{x_1 - x_2}\right) = \arg\left(\frac{x_3 - z}{x_1 - z}\right)$$

or

$$\arg\frac{(z - x_3)(x_1 - x_2)}{(z - x_1)(x_3 - x_2)} = 0$$

7.15. CIRCLES

A circle of radius r centred at z_0 is the locus of a point z at a constant distance r from z_0. Therefore the equation of this circle is

$$|z - z_0| = r.$$

Squaring it, we get

$$r^2 = |z - z_0|^2 = (z - z_0)(\bar{z} - \bar{z}_0) = z\bar{z} - z\bar{z}_0 - \bar{z}z_0 + z_0\bar{z}_0.$$

Alternatively, the equation of the circle is

$$|z|^2 = 2\,\mathrm{Re}(z\bar{z}_0) + r^2 - |z_0|^2.$$

7.15.1. *Example.* Appolonius of Perga (*circa* 260 – 190 BC) found that given two fixed points A_1 and A_2, the locus of a point P whose distances from A_1 and A_2 are in a constant ratio $1 : \mu$ (i.e. $A_2P = \mu A_1P$) is a circle.

For $\mu = 1/2$, Fig. 7.23 may suggest a geometric proof of the Appolonius' Theorem,

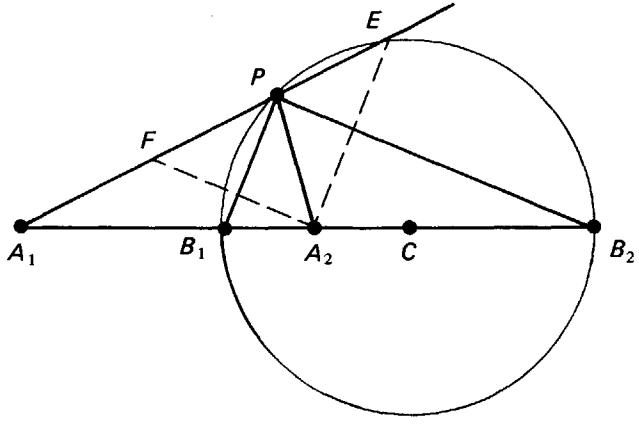


Fig. 7.23

where PB_1 and PB_2 are the internal and external bisectors of $\angle A_1PA_2$. Moreover, $FA_2 \parallel PB_2$ and $EA_2 \parallel PB_1$.

Let us prove the theorem using complex numbers. Choose the straight line passing through A_1 and A_2 as the real axis, and A_2 as the origin. Thus we have $A_1(a)$ and $A_2(0)$, where a is a real number. The condition on $P(z)$ is therefore simply the equation

$$\mu\,|z - a| = |z|$$

in the complex variable z. Squaring the left hand side, we obtain

$$\mu^2|z - a|^2 = \mu^2(z - a)(\bar{z} - a) = \mu^2(|z|^2 - 2\,\mathrm{Re}(za) + a^2)$$

Therefore the equation becomes

$$|z|^2(1 - \mu^2) = 2\,\mathrm{Re}(-\mu^2 az) + (\mu a)^2$$

$$\therefore \quad |z|^2 = 2\,\mathrm{Re}(\bar{z}_0 z) + r^2 - |z_0|^2$$

where

$$z_0 = \frac{-\mu^2 a}{1-\mu^2} \quad \text{and} \quad r = \left|\frac{\mu a}{1-\mu^2}\right|$$

proving that the point $P(z)$ describes a circle with radius $|\mu a/(1-\mu^2)|$ centred at $-\mu^2 a/(1-\mu^2)$ on the real axis.

7.16. STRAIGHT LINES

The simplest expression of a straight line on the plane is a parametric representation. Consider first the case where the straight line L passes through the origin O and another point $P(x_0)$ (Fig. 7.24).

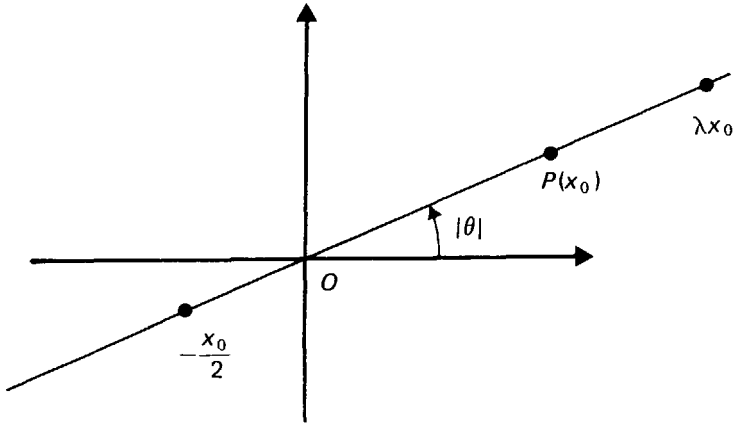


Fig. 7.24

Then every point on L is represented by a multiple λx_0 of x_0 where λ is a real number. Therefore the equation

$$z = \lambda x_0, \quad \lambda \in \mathbb{R}$$

is a *parametric representation* of the straight line L, in the sense that every point of L corresponds to a real number λ and vice versa.

If $\theta = \arg x_0$, then L is also the straight line that passes through O and has slope $\tan\theta$. Therefore, after absorbing $|x_0|$ into the parameter λ we rewrite $z = \lambda x_0$ as

$$z = \lambda(\cos\theta + i\sin\theta)$$

and take it to be the parametric representation of the straight line L with slope $\tan\theta$ and passing through O.

The next case to be considered is where the given straight line G has slope $\tan\theta$ and passes through a point $Q(x_1)$. G is parallel to L and can be obtained by a parallel translation of L in the direction of OQ (Fig. 7.25). Now a parallel translation in the plane corresponds to an addition of complex numbers. Therefore, the straight line G has the following parametric representation:

$$z = x_1 + \lambda x_0 \quad \text{or} \quad z = x_1 + \lambda(\cos\theta + i\sin\theta),$$

which is the so-called *point-slope form* in the parametric version.

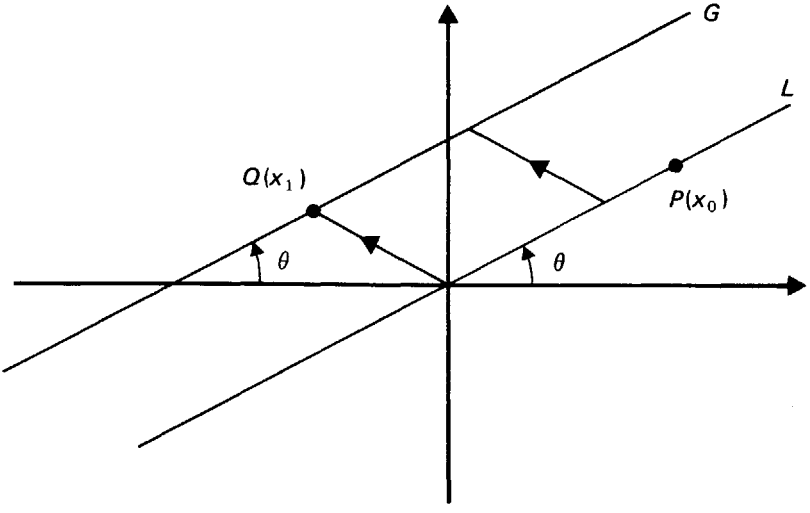


Fig. 7.25

Finally consider the straight line H that passes through two distinct points $P(x_0)$ and $Q(x_1)$. This straight line has slope $\tan\varphi$, where $\varphi = \arg(x_1 - x_0)$, and passes through $P(x_0)$ (Fig. 7.26). Therefore, H has the following parametric representation:

$$z = x_0 + \lambda(\cos\varphi + i\sin\varphi)$$

or

$$z = x_0 + \lambda(x_1 - x_0)$$

which may be regarded as the parametric version of the so-called *two-point form.*

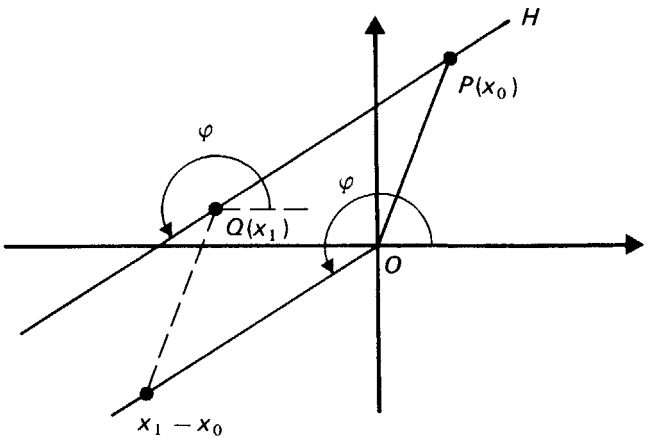


Fig. 7.26

7.16.1. *Example.* Let x_1, x_2, x_3 be three distinct points in the plane. Find the straight line N that passes through x_1 and is perpendicular to the straight line L that passes through x_2 and x_3 (Fig. 7.27).

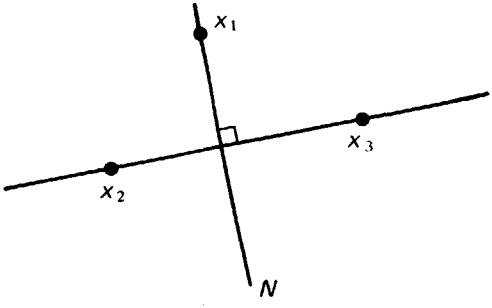


Fig. 7.27

Solution. The straight line L has slope $\tan\theta$, where $\theta = \arg(x_3 - x_2)$. Therefore, the perpendicular N has slope $\tan(\pi/2 + \theta)$, where

$$\frac{\pi}{2} + \theta = \arg i + \arg(x_3 - x_2) = \arg(x_3 i - x_2 i).$$

Using the point-slope form, we obtain

$$z = x_1 + \lambda i(x_3 - x_2).$$

Alternatively, a point z lies on N if and only if

$$\arg(z - x_1) = \arg(x_3 - x_2) \pm \frac{\pi}{2}$$

i.e.

$$\arg\left(\frac{z - x_1}{x_3 - x_2}\right) = \pm\frac{\pi}{2}$$

In other words, the complex number

$$\frac{z - x_1}{x_3 - x_2} = \frac{(z - x_1)(\overline{x}_3 - \overline{x}_2)}{|x_3 - x_2|^2}$$

is purely imaginary. Therefore N is represented by the equation

$$\mathrm{Re}(z - x_1)(\overline{x}_3 - \overline{x}_2) = 0$$

in the complex variable z. In contrast to the first solution, this is not a parametric representation. ■

The alternative solution to the problem of the example suggests that argument, real and imaginary parts, may be used to write the equation of a straight line.

7.16.2. *Example.* Show that

$$\mathrm{Im}\frac{z - x_1}{\cos\theta + i\sin\theta} = 0$$

is the equation of the straight line G which has slope $\tan\theta$ and passes through the point x_1 (Fig. 7.28).

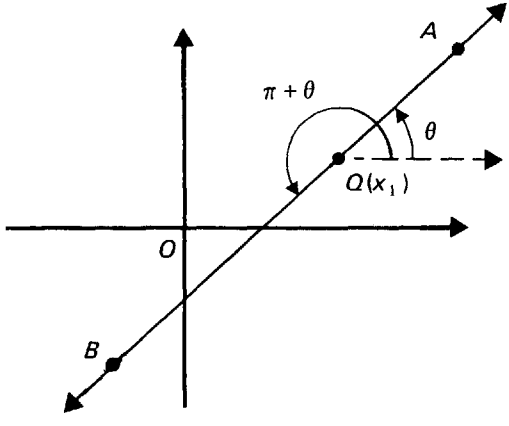


Fig. 7.28

Proof. The straight line G consists of two rays QA and QB, where A and B are two points of G lying on opposite sides of $Q(x_1)$. One of these rays makes an angle θ with the positive real axis and the other $\pi + \theta$. Therefore a point z falls on G if and only if $\arg(z - x_1)$ is θ or $\pi + \theta$. Now

$$\operatorname{Im} \frac{z - x_1}{\cos\theta + i\sin\theta} = 0 \quad \text{if and only if}$$

$$\arg \frac{z - x_1}{\cos\theta + i\sin\theta} = \arg(z - x_1) - \theta \quad \text{is 0 or } \pi,$$

i.e. $\arg(z - x_1)$ is θ or $\pi + \theta$. Therefore, the equation of G is simply

$$\operatorname{Im} \frac{z - x_1}{\cos\theta + i\sin\theta} = 0.$$

■

It follows easily that the two-point form of a straight line H that passes through x_0 and x_1 is given by

$$\operatorname{Im} \frac{z - x_1}{x_1 - x_0} = 0.$$

7.17. EXERCISE

1. Interpret the following locus in an Argand diagram:
$$|z + 3i|^2 - |z - 3i|^2 = 12.$$

2. If a and b are complex constants, interpret geometrically in an Argand diagram the following locus:
$$\arg\left(\frac{z - a}{z - b}\right) = \text{constant}.$$

3. If $|z_1| = |z_2| = |z_3|$ and $z_1 + z_2 + z_3 = 0$, prove that z_1, z_2 and z_3 are represented by the vertices of an equilateral triangle.

4. If a point $P(z)$ in an Argand diagram lies in the region above the real axis, prove that the number
$$\frac{z - i}{z + i}$$
is represented by a point inside the circle $|z| = 1$.

5. If z is a variable complex number such that $|z| = 1$ and $w = 2z + 1/z$, show that the point of the complex plane corresponding to w describes an ellipse.

6. If $|z| < 1$, prove that the principal value of $\arg(z + 1)$ lies between $-\pi/2$ and $\pi/2$.

7. When the vertices of a square $ABCD$ are taken anti-clockwise in that order, the points A and B represent the complex numbers $-1 + 4i$ and -3 respectively in an Argand diagram. Find the complex numbers represented by the other vertices and by the centre of the square.

8. If z is a complex number such that $|z + 1 + i| \leqslant 1$, find the maximum and minimum values of $|z|$.

9. Let a and b be real constants, and x, y, t be real variables. If
$$x + yi = a + \frac{b(1 + it)}{1 - it}$$
show that the locus of the point (x, y) as t varies is a circle.

10. Given that $3 + i$ and $4 + 2i$ are represented by the adjacent vertices of a square, find the possible numbers represented by the remaining vertices.

11. Prove that
    (i) if $|z_1 + z_2| = |z_1 - z_2|$, then the difference in the arguments of z_1 and z_2 is $\pi/2$;
    (ii) if $\arg\left(\frac{z_1 + z_2}{z_1 - z_2}\right) = \frac{\pi}{2}$, then $|z_1| = |z_2|$.

12. $PQRS$ is a parallelogram and X is the point of intersection of the diagonals. If P, R and S represent the numbers $1 + 3i$, $2 + 6i$ and $5 + 7i$ respectively, find the numbers represented by Q and X.

13. Triangles BCX, CAY and ABZ are described on the sides of a triangle ABC. If the points A, B, C, X, Y, Z in an Argand diagram represent the complex numbers a, b, c, x, y, z respectively and
$$\frac{x - c}{b - c} = \frac{y - a}{c - a} = \frac{z - b}{a - b},$$
show that the triangles BCX, CAY and ABZ are similar.

14. In an Argand diagram, points A and B represent the numbers $6i$ and 3 respectively. If a point P, which represents some complex number z, moves so that $PA = 2PB$, show that
$$z\bar{z} = (4 + 2i)z + (4 - 2i)\bar{z}.$$
Show that the locus of P is a circle and find its radius and the complex number corresponding to its centre.

15. If, in an Argand diagram, points representing the numbers $z_1, z_2, \ldots, z_n$ all lie on one side of some straight line through 0, prove that
$$\Sigma_{k=1}^{n} z_k \neq 0.$$

16. Complex numbers z_1 and z_2 are represented by points P_1 and P_2 respectively in an Argand diagram.
    (a) If $z_1(1 - z_2) = z_2$ and P_1 describes the line $2Re(z_1) + 1 = 0$, prove that P_2 describes a circle.
    (b) Find the sense in which the circle is described as P_1 moves along the line in the direction of increasing $\mathrm{Im}(z_1)$.

17. A complex number z is represented by a point P in an Argand diagram. If $(z - 1)/(z - i)$ is purely imaginary, prove that P moves on the circle with centre corresponding to $1/2 + i/2$ and radius $1/\sqrt{2}$.

18. Points P and Q represent the numbers z_1 and z_2 in an Argand diagram, and O is the origin. If $OP = OQ$, prove that
$$\frac{z_1 - z_2}{z_1 + z_2}$$
is of the form bi where b is real.

19. If points A, B, C representing the complex numbers z_1, z_2, z_3 in an Argand diagram are the vertices of an isosceles triangle, right-angled at B, prove that
$$(z_1 - z_2)^2 + (z_3 - z_2)^2 = 0.$$

20. Two complex numbers z and w are connected by the relation
$$w = (z - 1)(z + 1).$$
The point representing z in an Argand diagram describes the circle $|z| = 1$ in a counter-clockwise direction starting from $z = 1$. Find the path traced out by the point representing w.

21. By considering the modulus of the left-hand side of the following equation in z, or otherwise, prove that all the roots of
$$z^n \cos n\alpha + z^{n-1} \cos (n-1)\alpha + \ldots + z \cos \alpha = 1,$$
where α is real, correspond to points outside the circle $|z| = 1/2$ in an Argand diagram.

22. Given $w = z^2$, where $w = u + vi$ and $z = x + yi$ (u, v, x, y are real numbers),
    (a) Prove that $u = x^2 - y^2$
    and $v = 2xy$.
    (b) Prove that, when the point representing z describes the circle $|z| = 1$ in an Argand diagram, the point representing w describes the circle $|w| = 1$ twice.

23. Prove that the points representing the complex numbers $1, -1, a + bi, (a + bi)^{-1}$ (a, b are real numbers and not both equal to 0) in an Argand diagram are concyclic.

7.18. APPENDIX

We have seen that when we have extended the real number system $\mathbb{R}$ to the complex number system $\mathbb{C}$, as far as the provision of solutions of equations is concerned, there is no need to extend $\mathbb{C}$ any further. However we may still investigate the possibility and usefulness of extensions of $\mathbb{C}$ from another point of view. Since $\mathbb{R}$ is accepted as a 1-dimensional system and $\mathbb{C}$ a 2-dimensional system, we may now ask whether it is profitable to look for still higher dimensional number systems. The Irish mathematican, William Rowan Hamilton (1805 – 1865), and his German contemporary, Hermann Günther Grassmann (1809 – 1877), independently considered such problems and obtained results which are of a far-reaching character.

Hamilton found that, though theoretically possible, the 3-dimensional numbers, by necessity did not possess enough desirable properties to merit further attention. Going one dimension higher up, he then studied the 4-dimensional *number system* $\mathbb{H}$ *of quaternions*, which turned out to have many properties similar to those of $\mathbb{R}$ and $\mathbb{C}$.

The 4-dimensional numbers, which are called *quaternions*, are ordered quadruples (a, b, c, d) of real numbers. Sums and products of quaternions are defined as follows:

$$(a, b, c, d) + (e, f, g, h) = (a + e, b + f, c + g, d + h)$$

$$(a, b, c, d)(e, f, g, h)$$
$$= (ae - bf - cg - dh, af + be + ch - dg, ag - bh + ce + df,$$
$$ah + bg - cf + de).$$

Similar to what was done in Section 7.5, after replacing each quaternion of the form $(a, 0, 0, 0)$ by the real number a and each quaternion of the form $(a, b, 0, 0)$ by the complex number $a + bi$, we may think of both $\mathbb{R}$ and $\mathbb{C}$ as subsystems of the number system $\mathbb{H}$. The next task to perform is the verification of the usual rules of arithmetic. It turns out that with the single exception of the commutative law of multiplication, all the usual rules hold true for quaternions in exactly the same manner as they hold true for complex numbers. That the multiplication of quaternions is non-commutative can be seen from the following examples. Following the definition of product, we have

$$(0, 1, 0, 0)(0, 0, 1, 0) = (0, 0, 0, 1)$$
$$(0, 0, 1, 0)(0, 1, 0, 0) = (0, 0, 0, -1).$$

Therefore

$$(0, 1, 0, 0)(0, 0, 1, 0) \neq (0, 0, 1, 0)(0, 1, 0, 0).$$

Similiar to the standard notation of complex numbers by which each complex number is written in terms of the real unit 1 and the imaginary unit i, quaternions may be expressed in terms of the four *quaternion units*

$$1 = (1, 0, 0, 0); \quad i = (0, 1, 0, 0)$$
$$j = (0, 0, 1, 0); \quad k = (0, 0, 0, 1)$$

Thus

$$(a, b, c, d) = a + bi + cj + dk.$$

By definition, the various products of the quaternion units are

$$i^2 = j^2 = k^2 = -1$$
$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j$$

which may be tabulated as follows.

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Historically, the system $\mathbb{H}$ of quaternions is the first non-commutative number system ever discovered.

Independently of Hamilton, Grassmann condidered ordered n-tuples

$(a_1, a_2, \ldots, a_n)$ of real numbers. These are called *hypercomplex numbers.* Again similar to the standard notation of complex numbers, these may be written in terms of n fundamental units $e_1, e_2, \ldots, e_n$. Thus

$$(a_1, a_2, \ldots, a_n) = a_1e_1 + a_2e_2 + \ldots + a_ne_n = \Sigma\, a_ie_i.$$

For these n-dimensional numbers, the sum is defined by

$$\Sigma\, a_ie_i + \Sigma\, b_ie_i = \Sigma\, (a_i + b_i)e_i.$$

It follows easily from the definition that the addition of hypercomplex numbers satisfy all the usual rules similar to those for the complex numbers. The product of two hypercomplex numbers is given by

$$(\sum_i a_ie_i)\,(\sum_j b_je_j) = \sum_{i,j} a_ib_je_ie_j$$

where the products e_ie_j of the fundamental units are to be found in a multiplication table similar in form to that of the quaternions above. Different multiplication tables (even with the same dimension n) will give rise to different types of hypercomplex numbers and different types of hypercomplex numbers may satisfy different rules of arithmetic. Each type of hypercomplex number is called a *Grassmann algebra*. It is, therefore, entirely feasible that some Grassmann algebras may not obey some of the usual rules of multiplication. This means that they may be non-commutative or non-associative or both.

Finally we remark that the complex number system $\mathbb{C}$, the system $\mathbb{H}$ of quaternions, and every possible Grassmann algebra are much more than just games of symbols, created by the whimsicality of mathematicians. They are constructed to solve mathematical problems which are relevant to our physical world. Complex numbers are found to be indispensable tools of classical physics and engineering. Quaternions and hypercomplex numbers are just as useful in modern physics and electronic engineering.

ANSWERS TO EXERCISES

1.5

1. All
2. (a), (c) and (d)
3. $F \subset A$, $B \subset D$, $E \subset C$
4. (a), (d)
5. $A = \phi$, $B = \{\phi\}$ and $C = \{\phi, \{\phi\}\}$
6. ϕ, $\{\phi\}$, $\{\phi, \{\phi\}\}$, $\{\phi, \{\phi\}, \{\phi, \{\phi\}\}\}$
7. $\{a, b, c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, $\{a\}$, $\{b\}$, $\{c\}$, ϕ
8. None
9. A, $\{0\}$, $\{\{1, 2\}\}$, ϕ
10. All

1.9

1. (a) $\{a, b, c, d, e\}$, $\{a, c, d, e\}$, $\{a, b, c, d, e\}$
   (b) $\{c, d\}$, $\{e\}$, $\{a\}$
2. (a) C, E
   (b) D, E
   (c) A, B, D
   (d) None
3. (b), (c), (d)
4. (a), (b), (c)
5. Not always true; consider $a = c \neq b$
7. $A = \{0, 1\}$, $B = \{1\}$ and $C = \{0\}$
13. All are false
16. (b) $\{\{1, 2\}, \{2, 3, 4\}, \{4\}\}$
17. (a) neither
    (b) subset
    (c) element
    (d) subset and element
    (e) element
24. (b) $(A \mid A) \mid (B \mid B)$
25. (a) $A_1 \cap A_2 \cap \ldots \cap A_i$
    (b) $T_1 = A_1$, $T_i = A_i \setminus (A_1 \cup A_2 \cup \ldots \cup A_{i-1})$ $(i = 2, 3, \ldots, n)$

1.11

1. $\{(1, a), (2, a), (3, a), (1, b), (2, b), (3, b)\}$,
   $\{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}$
3. (a, a)
4. $A = C = \{1\}$, $B = D = \{2\}$

1.14
1. (a) Yes
   (b) No
   (c) No
2. (a) Yes
   (b) Yes

10. No; no

20. (b) No

2.4
1. (a) Yes; -1
   (b) Yes; 2
   (c) No

2.9
6. (a) $\dfrac{n(n+1)}{2}$
   (b) $\dfrac{a}{(1-a)^2}$, $\dfrac{1}{1-a}$, $-\dfrac{a}{(1-a)^2}$

3.2
1. 90
2. 24
3. 2
4. 5880; 3000
5. $(2n-1) \times (2n-3) \times \ldots \times 1$

3.5
1. 360
2. 2880
3. 3456
4. 1440
5. $(n-1)!\,(n-2)$
6. $(n-2)!\,(n^2-3n+3)$
7. $(n-2)!\,(n-3)(n+2)$

3.7
1. 125
2. 1024
3. (a) 648
   (b) 320
4. 256; 232
5. $\dfrac{n(n^r-1)}{n-1}$

7. 6198

8. (a) $2^n - 2^{\frac{n}{2}}$ if n is even;

$2^n - 2^{\frac{n+1}{2}}$ if n is odd

(b) $\frac{n}{2} \times 2^{\frac{n}{2}}$ if n is even;

$\frac{n-1}{2} \times 2^{\frac{n+1}{2}}$ if n is odd

3.10

1. 126
2. (a) 700
   (b) 340
   (c) 174
3. $\frac{(m+1)!}{n!(m+1-n)!}$
4. (a) 144
   (b) 12
5. 144
6. 360
7. 30
8. 3600

3.12

1. 35
2. (i) 84
   (ii) 84
3. (i) 330
   (ii) 462
4. (i) 2002
   (ii) 6
   (iii) 1876
5. 2024
6. $\frac{(mn)!}{(m!)^n\, n!}$
7. (a) 225
   (b) 465
   (c) 240

8. 26; 52
9. (a) 10
   (b) 10
   (c) 45
   (d) 6
10. 5760
11. 2520
12. 2366
13. $\dfrac{(m+2)(m+1)(n+2)(n+1)}{4}$
14. 502
16. $\dfrac{n!}{r!\,s!\,(n-r-s)!}$
17. $\dfrac{n!}{a!\,b!\,c!}$
18. $\dfrac{n(n-4)(n-5)}{6}$
19. $\dfrac{(n-3)(n^2-9n+26)}{6}$
21. 36

3.14
1. $2^n - 1$
2. 300
3. 875
4. 15
8. 2001

3.16
6. (a) $\dfrac{n(n+1)}{2}$
13. (a) $x^4 + 4x^2 + 6 + \dfrac{4}{x} + \dfrac{1}{x^4}$
    (b) $32x^5 - 400x^4 + 2000x^3 - 10000x^2 + 6250x - 3125$
14. -10
15. $\pm\sqrt{\dfrac{14}{3}}$
16. $n = 16, a = 8$
17. $1 + 8x + 36x^2 + 104x^3 + \ldots$
18. $-2; -345$

19. -576
21. $a = \pm 2, b = \pm 1$
22. $m = 3, n = 2$
23. (b) (i) $(xy)^2 - xy - 420 = 0$
    (iii) $x = 5, y = -4$
24. $x = 3, a = 2, n = 6$
25. $y^7 - 7y^5 + 14y^3 - 7y$
26. $a = 7, b = 14, c = 7$
27. (a) $3 - 3x + 9x^2 - 15x^3 + \ldots$
    (b) $-\frac{1}{2} < x < \frac{1}{2}$

4.2
2. $-3, -4, -5$
3. $0, -1, -2, -3, -4, -5$

4.5
1. The converse is not necessarily true; consider $a_1 = a_2 = b = 3$
2. 24
3. 60
6. (i) $q = 0, r = 1$
   (ii) $q = -1, r = 1$
   (iii) $q = -2, r = 4$
   (iv) $q = 1, r = 3$

4.7
15. (b) R.H.S. is 2 x gcd(m, n) if and only if m and n are odd integers

4.10
7. (b) 96, 120
8. (b) True; false
9. 17
10. 22338
13. $m = 4, n = -3$

4.14
7. No
12. (a) No
22. $\prod_{i=1}^{r} (a_i + 1)$

4.17
3. 8
4. 37
9. $x \equiv 4, y \equiv 3$

5.2

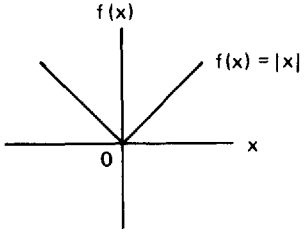
4. None; the series is divergent
9. (a) Yes; not necessarily
   (b) No
   (c) Yes
   (d) Yes

5.4

3. (i) Incorrect
   (ii) Correct
   (iii) Incorrect
   (iv) Incorrect
4. 1.414214; 2.236068; 1.571429

5.6

28.



5.9

5. No
10. sup $A = C$
11. (b) No
12. No

5.13

1. $x = 5, y = -2$
2. $\frac{1}{4}$
5. $\frac{15}{4}$
7. $x = 2, y = 8$
8. 3
9. 2
11. $\begin{cases} x = 2 \\ y = \frac{1}{8} \end{cases}$ or $\begin{cases} x = 64 \\ y = 4 \end{cases}$

6.2 1. 0
7. 0

6.5 1. 3
3. (c) 2

6.16 1. 0
5. $\frac{1}{4}\left[\frac{1}{3} - \frac{1}{(2n+1)(2n+3)}\right]$; $\frac{1}{12}$
6. $\frac{1}{2}\left[\frac{1}{6} - \frac{1}{(n+2)(n+3)}\right] + \frac{2}{3}\left[\frac{1}{6} - \frac{1}{(n+1)(n+2)(n+3)}\right]$; $\frac{7}{36}$
11. (a) $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$
(b) $\frac{1}{4}\left[3^n \sin\frac{\theta}{3^n} - \sin\theta\right]$; $\frac{1}{4}(\theta - \sin\theta)$

7.6 1. $34 + 22i$
2. $2 + 7i$
3. $2i$
4. $2 + 11i$
5. $-2 - 16i$
6. -10
7. $-i$
8. $x^2 + y^2$
9. $\frac{3}{25} + \frac{4}{25}i$
10. i
11. $\frac{5}{7} - \frac{6}{7}i$
12. $\frac{1}{2} - \frac{1}{2}i\tan\frac{\theta}{2}$
13. $x = 1, y = 2$
14. $x = \frac{2}{29}, y = -\frac{5}{29}$
15. $x = -4, y = 0$
16. $p = 7, q = 3$
18. $\frac{(c^2 - a^2)^2 + a^2(b+c)^2}{(c-b)^2}$

19. $2;\ \frac{2\pi}{3}$

21. $\frac{1}{2} \pm \frac{\sqrt{3}}{2} i$

22. (i) $\tan \frac{\theta}{2}\ ; -\frac{\pi}{2}$

    (ii) $-\tan \frac{\theta}{2}\ ;\ \frac{\pi}{2}$

7.9

7. $-2 \pm 3i$
8. $\pm \frac{3}{2} i$
9. $\cos \theta \pm i \sin \theta$
10. $1 + 4i$ or $-\frac{1}{3} - \frac{1}{3} i$
11. $x^2 - 4x + 7 = 0$
12. $d^2 + 4b^2 c = 4abd$

7.13

1. $\pm\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\, i\right)$
2. $\pm(3 + 2i)$
3. $\frac{1 \pm \sqrt{3}\, i}{2}\ , -1$
4. $\sqrt[3]{2 \sin \frac{\theta}{2}}\ \left[\cos \frac{(4k-1)\pi + \theta}{6} + i \sin \frac{(4k-1)\pi + \theta}{6}\right]$
   $(k = 0, 1, 2)$
5. $(x^2 - \sqrt{3}x + 1)(x^2 + \sqrt{3}x + 1)(x^2 + 1)$
6. (a) $\cos \frac{2k+1}{7} \pi + i \sin \frac{2k+1}{7} \pi \quad (k = 0, 1, 2, \ldots, 6)$

10. (a) $-\frac{1}{2}\left(1 \pm i \cot \frac{k\pi}{8}\right) \quad (k = 1, 2, 3)$

11. $\frac{a}{2}\left(1 + i \cot \frac{2k+1}{2n} \pi\right) \quad (k = 0, 1, 2, \ldots, n-1)$

12. 19

14. (a) $\pm i \tan \frac{2k+1}{4n} \pi \quad (k = 0, 1, 2, \ldots, n-1)$

7.17

1. The line $\mathrm{Im} Z = 1$
2. The major arc AB or the minor arc AB of a circle through A and B.
7. C: $1 - 2i$, D: $3 + 2i$ and centre: i
8. $\sqrt{2} + 1; \sqrt{2} - 1$
10. $\begin{cases} 3 + 3i \\ 2 + 2i \end{cases}$ or $\begin{cases} 5 + i \\ 4 \end{cases}$
12. $-2 + 2i; \frac{3}{2} + \frac{9}{2}i$
14. $2\sqrt{5}; 4 - 2i$
16. (b) Clockwise
20. W moves along the imaginary axis from the origin to ∞ and returns to the origin from $-\infty$.

INDEX